

GenCore version 4.5						
Copyright (c) 1993 - 2000 Compugen Ltd.						
MM Protein - protein search, using sw model						
run on:	September 17, 2002, 15:42:29	Search time 12.89 Seconds				
sequence:	1 MKSTAIITIVLVAFCTILEDG.....	(without alignments)				
scoring table:	BLOSUM62	IRRGGFCRGTERTTCVCYR 61				
perfect score:	339					
title:	US-09-829-481-4					
gapext:	0.5					
gapopen:	10.0					
total number of hits satisfying chosen parameters:	231628					
minimum DB seq length:	0					
maximum DB seq length:	2000000000					
post-processing:	Maximum Match 0%					
	Listing first 45 summaries					
database :	Issued_Patents_AA.*					
	1: /cgn2_6/ptodata/2/1aa/5A..COMB.pep:*					
	2: /cgn2_6/ptodata/2/1aa/5B..COMB.pep:*					
	3: /cgn2_6/ptodata/2/1aa/5C..COMB.pep:*					
	4: /cgn2_6/ptodata/2/1aa/6A..COMB..pep:*					
	5: /cgn2_6/ptodata/2/1aa/6B..COMB..pep:*					
	6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*					
searched:	231628 seqs	24425594 residues				
SUMMARIES						
result No.	Score	Query Length	DB ID	Description		
1	68.5	20.2	72	4 US-09-003-198A 17		
2	66	19.5	40	1 US-08-385-175-1	Sequence 17, Appli	
3	66	19.5	40	1 US-08-385-175-39	Sequence 16, Appli	
4	64.5	19.0	80	4 US-09-103-375-39	Sequence 39, Appli	
5	64	18.9	74	4 US-09-442-031-4	Sequence 20, Appli	
6	63.5	18.7	768	1 US-08-454-455-4	Sequence 4, Appli	
7	63.5	18.7	769	1 US-08-454-455-6	Sequence 6, Appli	
8	61.5	18.1	80	1 US-08-377-087-49	Sequence 49, Appli	
9	61.5	18.1	80	2 US-08-777-192-49	Sequence 49, Appli	
10	61.5	18.1	80	4 US-08-971-982-49	Sequence 49, Appli	
11	60.5	17.8	81	4 US-09-053-021-4	Sequence 4, Appli	
12	60.5	17.8	81	4 US-09-053-021-9	Sequence 9, Appli	
13	60.5	17.8	652	2 US-08-751-305-2	Sequence 2, Appli	
14	60	17.7	79	1 US-08-627-706-15	Sequence 15, Appli	
15	60	17.7	79	4 US-09-103-889-15	Sequence 15, Appli	
16	59.5	17.6	42	2 US-08-751-305-3	Sequence 3, Appli	
17	59.5	17.6	80	1 US-08-377-687-59	Sequence 59, Appli	
18	59.5	17.6	80	2 US-08-777-192-59	Sequence 59, Appli	
19	59.5	17.6	80	4 US-08-971-982-59	Sequence 59, Appli	
20	59	17.4	204	1 US-08-652-859-2	Sequence 2, Appli	
21	59	17.4	204	2 US-08-919-706-2	Sequence 2, Appli	
22	59	17.4	204	2 US-09-153-51-2	Sequence 2, Appli	
23	58.5	17.3	75	2 US-08-289-458-2	Sequence 2, Appli	
24	58.5	17.3	75	2 US-08-761-549-2	Sequence 2, Appli	
25	58.5	17.3	75	4 US-09-127-646-2	Sequence 2, Appli	
26	57.5	17.0	221	2 US-08-925-701-1	Sequence 1, Appli	
27	57	16.8	86	1 US-08-149-839B 14	Sequence 14, Appli	

Db 4 KSLACLSFLILVLVLFVAQEIVVSEANTCENLAGSYKGVCFGCDRHCR - QEGATSGR CR 61
 Qy 51 GTFRITC 57
 Db 62 DDFRCWC 68

RESULT 2
 US-08-385-375-16
 Sequence 16, Application US/08385375
 Patent No. 5631144
 GENERAL INFORMATION:
 APPLICANT: LEMOINE, Yves
 APPLICANT: NGUYEN, Martine
 APPLICANT: ACHSTETTER, Tilman
 APPLICANT: REICHHART, Jean-Marc
 APPLICANT: SWECKER & Mathis
 ADDRESS: Burns, Doane, Swecker & Mathis
 STREET: P.O. Box 1404
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: United States
 ZIP: 22313-1404
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patientin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 NUMBER OF SEQUENCES: 40
 CORRESPONDENCE ADDRESS:
 STREET: Burns, Doane, Swecker & Mathis
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: United States
 ZIP: 22313-1404
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patientin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/385,375
 FILING DATE: 2000-08-08
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/178,356
 FILING DATE: 04-JAN-1994
 APPLICATION NUMBER: FR 89/05687
 FILING DATE: 28-APR-1999
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/FR90/00306
 FILING DATE: 27-APR-1990
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/178,356
 FILING DATE: 04-JAN-1994
 APPLICATION NUMBER: FR 89/05687
 FILING DATE: 28-APR-1999
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/FR90/00306
 FILING DATE: 27-APR-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: crane-peury, Sharon E
 REGISTRATION NUMBER: 36,113
 REFERENCE/DOCKET NUMBER: 017753-009
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 FAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 40 amino acids
 STRANDEDNESS: single
 TOPOLogy: linear
 MOLECULE TYPE: peptide
 US-08-385-375-16

Query Match 19.5%; Score 66; DB 1; Length 40;
 Best Local Similarity 39.4%; Pred. No. 0.28; Pred. No. 0.28;
 Matches 13; Conservative 2; Mismatches 16; Indels 2; Gaps 1;
 Qy 27 GCPFNAGKCHRHKCKSIRRGGFCRGTFRTTCYC 59
 Db 8 GTGINHSACAACHLRLRGNGGYCNG - KGVCYC 38

RESULT 3
 Sequence 39, Application US/08385375
 Patent No. 5631144
 GENERAL INFORMATION:
 APPLICANT: LEMOINE, Yves
 APPLICANT: NGUYEN, Martine
 APPLICANT: ACHSTETTER, Tilman
 APPLICANT: REICHHART, Jean-Marc
 APPLICANT: SWECKER & Mathis
 ADDRESS: Burns, Doane, Swecker & Mathis
 STREET: P.O. Box 1404
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: United States
 ZIP: 22313-1404
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patientin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/385,375
 FILING DATE: 2000-08-08
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/178,356
 FILING DATE: 04-JAN-1994
 APPLICATION NUMBER: FR 89/05687
 FILING DATE: 28-APR-1999
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/FR90/00306
 FILING DATE: 27-APR-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: crane-peury, Sharon E
 REGISTRATION NUMBER: 36,113
 REFERENCE/DOCKET NUMBER: 017753-009
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 FAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 39:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 40 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLogy: linear
 MOLECULE TYPE: protein
 US-08-385-375-39

Query Match 19.5%; Score 66; DB 1; Length 40;
 Best Local Similarity 39.4%; Pred. No. 0.28; Pred. No. 0.28;
 Matches 13; Conservative 2; Mismatches 16; Indels 2; Gaps 1;
 Qy 27 GCPFNAGKCHRHKCKSIRRGGFCRGTFRTTCYC 59
 Db 8 GTGINHSACAACHLRLRGNGGYCNG - KGVCYC 38

RESULT 4
 Sequence 20, Application US/09103489
 Patent No. 6215048
 GENERAL INFORMATION:
 APPLICANT: Liang, Jihong
 APPLICANT: Shah, Dilip M.
 APPLICANT: Wu, Yonnie S.

APPLICANT: Rosenberger, Cindy A.
 TITLE OF INVENTION: Antifungal Polypeptide and Methods for Controlling Plant Pathogenic Fungi
 NUMBER OF SEQUENCES: 200
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Charles E. Cohen, Monsanto Company, BB4F STREET: 700 Chesterfield Village Parkway No. 6215048th STATE: St. Louis COUNTRY: USA ZIP: 63198
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/103,489
 FILING DATE: 24-JUN-1998
 CLASIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Cohen, Charles E.
 REGISTRATION NUMBER: 34,565
 REFERENCE/DOCKET NUMBER: 38-21 (10700)A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (314) 537-6224
 TELEFAX: (314) 537-6047
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 80 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-09-103-489-20

Query Match 19 0%; Score 64.5; DB 4; Length 80;
 Best Local Similarity 30 0%; Pred. No. 0.84%; Indels 13; Gaps 4;
 Matches 21; Conservative 9; Mismatches 27;

Qy 4 IAIIFIVLVAFCILE-DGIVEAGPGCPNAG-----KCHRHKSIIRR-RGGPCRQ 51
 Db 8 IAIFFAALFLFAAFAETINVEAQKLCRESGTWSGVCGNNACKNQCTNLERKHGSQN 67

Query Match 5 0%; Score 64.5; DB 4; Length 80;
 Best Local Similarity 30 0%; Pred. No. 0.84%; Indels 13; Gaps 4;
 Matches 21; Conservative 9; Mismatches 27;

Qy 52 TFRIT-TCVCY 60
 Db 68 VFPAAHKC1CY 77

RESULT 5
 US-09-442-631-4
 Sequence 4, Application US/09442631
 Patent No. 6300489
 GENERAL INFORMATION:
 APPLICANT: OH, BOUNG-JUN
 APPLICANT: KO, MOON KYUNG
 APPLICANT: SHIN, BYONCHUL
 APPLICANT: CHUNG, CHANG HO
 TITLE OF INVENTION: SMALL AND CYSTEINE RICH ANTIFUNGAL DEFENSIN AND THIONIN-LIKE PROTEIN GENES HIGHLY EXPRESSED IN THE FILE REFERENCE: 1942/44
 CURRENT APPLICATION NUMBER: US/09/442,631
 CURRENT FILING DATE: 1999-11-18
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 4
 LENGTH: 74
 TYPE: PRT
 ORGANISM: Capsicum annuum
 US-09-442-631-4

Query Match 18 9%; Score 64; DB 4; Length 74;
 Best Local Similarity 35 8%; Pred. No. 0.88%; Indels 12; Gaps 5;
 Matches 24; Conservative 8; Mismatches 23;

Qy 2 KSIAILIFIVLVAFC--IILEDGIVEAGPGCPFN-AGKC--HRHKSIIRR-RGGPCRQ 52
 Db 6 KVVATIFLMMKVFATDMAEAKICAEALSG --NPKGLCLSSRDCGNVRREGFTDSCIG 62

Query Match 53 FRTTCVCY 59
 Db 63 FRLQCFC 69

RESULT 6
 US-08-454-455-4
 Sequence 4, Application US/08454455
 Patent No. 56335601
 GENERAL INFORMATION:
 APPLICANT: Moyle, Matthew
 APPLICANT: McLean, John W.
 TITLE OF INVENTION: NOVEL BETA INTEGRIN SUBUNIT
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 460 Point San Bruno Blvd
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WinPatin (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/454,455
 FILING DATE: 30-MAY-1995
 CLASSIFICATION: 530
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 08/193989
 FILING DATE: 09-FEB-1994
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 08/004142
 FILING DATE: 13-JAN-1993
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 07/670607
 FILING DATE: 14-MAR-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Lee, Wendy M.
 REGISTRATION NUMBER: 00,000
 REFERENCE/DOCKET NUMBER: P0699C2D2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/225-1994
 TELEFAX: 415/952-9881
 TELEX: 910/371-7168
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 768 amino acids
 TYPE: Amino Acid
 TOPOLOGY: Linear
 US-08-454-455-4

Query Match 18 7%; Score 63.5; DB 1; Length 768;
 Best Local Similarity 24 7%; Pred. No. 10; Indels 39; Gaps 5;
 Matches 19; Conservative 8; Mismatches 11;

Qy 14 FCILEDGIVEAGPGCPN-----AGKCH-----RHKSIRR 44
 Db 545 YCEKDD-----FSCPIFHGSILCAGHGEAQRCCQPSAAQHC -VNS 596
 Qy 45 RGGFRCGTFRITCVCYR 61

RESULT 7
 US-08-454-455-6
 Sequence 6, Application US/08454455
 ; Patent No. 5635601
 ; GENERAL INFORMATION:
 ; APPLICANT: MOYLE, Matthew W.
 ; TITLE OF INVENTION: NOVEL BETA INTEGRIN SUBUNIT
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Gentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Winpatin (Genetech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/454,455
 ; FILING DATE: 30-MAY-1995
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/193989
 ; FILING DATE: 09-FEB-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/004142
 ; FILING DATE: 13-JAN-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/670607
 ; FILING DATE: 14-MAR-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lee, Wendy M.
 ; REGISTRATION NUMBER: 00-000
 ; REFERENCE/DOCKET NUMBER: P0699C2D2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-1994
 ; TELEFAX: 415/952-9881
 ; TELEFAX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 769 amino acids
 ; TYPE: Amino Acid
 ; TOPOLOGY: Linear
 ; US-08-454-455-6

Query Match 18.7%; Score 63.5%; DB 1; Length 769;
 Best Local Similarity 24.7%; Pred. No. 10; Indels 39; Gaps 5;

RESULT 8
 US-08-377-687-49
 ; Sequence 49, Application US/08377687
 ; Patent No. 558525
 ; GENERAL INFORMATION:
 ; APPLICANT: BROEKERT, WILLEM F.

Query Match 18.7%; Score 63.5%; DB 1; Length 769;
 Best Local Similarity 24.7%; Pred. No. 10; Indels 39; Gaps 5;

RESULT 9
 US-08-777-192-49
 ; Sequence 49, Application US/08777192
 ; Patent No. 5824869
 ; GENERAL INFORMATION:
 ; APPLICANT: BROEKERT, WILLEM F.
 ; APPLICANT: CAMMUE, BRUNO P.A.
 ; APPLICANT: OSBORN, RUPERT W.
 ; APPLICANT: REES, SARAH B.
 ; APPLICANT: TERRAS, FRANKY R.G.
 ; APPLICANT: VANDERLEDEN, JOZEFF
 ; TITLE OF INVENTION: BIOCIDAL PROTEINS
 ; NUMBER OF SEQUENCES: 59
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
 ; STREET: 1100 NEW YORK AVENUE, N.W.
 ; CITY: WASHINGTON
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/377,687
 ; FILING DATE:
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/002,480
 ; FILING DATE: 04-JAN-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: KOURILIS, PAUL N.
 ; REGISTRATION NUMBER: 16,773
 ; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-861-3000
 ; TELEFAX: 202-822-0944
 ; INFORMATION FOR SEQ ID NO: 49:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 90 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-377-687-49

Query Match 18.1%; Score 61.5%; DB 1; Length 80;
 Best Local Similarity 30.0%; Prod. No. 1,9;
 Matches 21; Conservative 8; Mismatches 28; Indels 13; Gaps 4;

Qy 4 IAIIFTIVLVAFCILE-PGIVEAGFGCFNAG-----RCKHSIRR-RGGPCRG 51
 Db 8 IAIFFEALVLFAAFEAFTMVEAKLICERPSTWSGVCGNNNACKNQCINLEKARHGSCNY 67

Qy 52 TFRF-TFCVY 60
 Db 68 VFAHKRCIY 77

Query Match 18.7%; Score 63.5%; DB 1; Length 769;
 Best Local Similarity 24.7%; Pred. No. 10; Indels 39; Gaps 5;

RESULT 10
 US-08-454-455-6
 ; Sequence 45, Application US/08454455
 ; Patent No. 5635601
 ; GENERAL INFORMATION:
 ; APPLICANT: VANDERLEDEN, JOZEFF
 ; TITLE OF INVENTION: BIOCIDAL PROTEINS
 ; NUMBER OF SEQUENCES: 59
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
 ; STREET: 1100 NEW YORK AVENUE, N.W.
 ; CITY: WASHINGTON
 ; STATE: D.C.

COUNTRY: USA
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/777,192
 FILING DATE: 04-JAN-1993
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/002,480
 FILING DATE: 04-JAN-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: KOKULIS, PAUL N.
 REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-861-3000
 TELEFAX: 202-822-0944
 INFORMATION FOR SEQ ID NO: 49:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 80 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-777-192-49

Query Match 18.1%; Score 61.5; DB 2; Length 80;
 Best Local Similarity 30.0%; Pred. No. 1.9;
 Matches 21; Conservative 8; Missmatches 28; Indels 13; Gaps 4;
 RESULT 11

Qy 4 IAIIFIVLVAFCILE-DGIVEAEGFGCPFNAG-----KCCHRICKSIRR-RGGFCRG 51
 ||::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 8 IALLFAALVLFAAFEAPTMVEAQKLCLERSGTWSGVCGNNNACKNQCNLEKARIHGSCNY 67
 - Qy 52 TFRP-TCYCY 60
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 68 VFPAHKCICY 77
 RESULT 10

Sequence 49 Application US/08971982
 Pat. No. 6187304
 GENERAL INFORMATION:
 APPLICANT: BROekaert, Willem F.,
 CAMMUE, Bruno P.A.,
 OSSBORN, Rupert W.,
 REES, Sarah B.,
 TERRAS, Franky R.G.
 VANDERLYDEN, JOSEPH
 TITLE OF INVENTION: BIOTIDAL PROTEINS
 NUMBER OF SEQUENCES: 59
 CORRESPONDENCE ADDRESS:
 ADDRESSSEE: CUSHMAN DARBY & CUSHMAN
 STREET: 1100 NEW YORK AVENUE, N.W.
 CITY: WASHINGTON
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/971,982
 FILING DATE: 17-No. 6187304-1997
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/002,480
 FILING DATE: 04-JAN-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: KOKULIS, PAUL N.
 REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-861-3000
 TELEFAX: 202-822-0944
 INFORMATION FOR SEQ ID NO: 49:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 80 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-971-982-49

Query Match 18.1%; Score 61.5; DB 4; Length 80;
 Best Local Similarity 30.0%; Pred. No. 1.9;
 Matches 21; Conservative 8; Missmatches 28; Indels 13; Gaps 4;
 RESULT 11

Qy 4 IAIIFIVLVAFCILE-DGIVEAEGFGCPFNAG-----KCCHRICKSIRR-RGGFCRG 51
 ||::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 8 IALLFAALVLFAAFEAPTMVEAQKLCLERSGTWSGVCGNNNACKNQCNLEKARIHGSCNY 67
 - Qy 52 TFRP-TCYCY 60
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 68 VFPAHKCICY 77
 RESULT 11

Sequence 49 Application US/09053021
 Pat. No. 6220785
 GENERAL INFORMATION:
 APPLICANT: SELSKO, Barbara
 GARCIA-RODRIGUEZ, Consuelo
 APPLICANT: ZAMUDIO-ZUNIGA, Fernando
 BECERRIL-LUJAN, Baltazar
 APPLICANT: POSSANI-POSTAY, Lourival D.
 APPLICANT: POSSANI-POSTAY, Lourival D.
 TITLE OF INVENTION: Primary Sequence and cDNA of
 Patent No. 6220785
 TITLE OF INVENTION: Insecticidally Effective Toxins from Scorpions of the
 Genus Centruroides
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Finnean, Henderson, Farabow, Garrett &
 Dunner, L.L.P.
 STREET: 1300 I Street, N.W.
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/053,021
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/017,007
 FILING DATE: 30-APR-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/848,261
 FILING DATE: 29-APR-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Garrett, Arthur S.
 REGISTRATION NUMBER: 20,338

; REFERENCE/DOCKET NUMBER: 06899.0001-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 4;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 81 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-053-021-4

; LENGTH: 81 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-053-021-9

Query Match 17.8%; Score 60.5; DB 4; Length 81;
Best Local Similarity 23.4%; Pred. No. 2, 4;
Matches 15; Conservative 10; Mismatches 28; Indels 11; Gaps 3;

Query 4 IAIIFIVLVAFCLIEDGIVEAGFGCPFNA-----GKCHRICKS1TRRG--GFCRGTFRT 55
Db 1 ITACQLVLTGCAKEGYLNKSYNCLLGENKNCMDMECKAKNQGSYGC--YKL 57

Query Match 17.8%; Score 60.5; DB 4; Length 81;
Best Local Similarity 23.4%; Pred. No. 2, 4;
Matches 15; Conservative 10; Mismatches 28; Indels 11; Gaps 3;

Query 4 IAIIFIVLVAFCLIEDGIVEAGFGCPFNA-----GKCHRICKS1TRRG--GFCRGTFRT 55
Db 1 ITACQLVLTGCAKEGYLNKSYNCLLGENKNCMDMECKAKNQGSYGC--YKL 57

RESULT 13
US-08-751-305-2
Sequence 2, Application US/08751305
; Patent No. 5965439
; GENERAL INFORMATION:
; APPLICANT: Tenner et al., Andrea J.
; TITLE OF INVENTION: HOST DEFENSE ENHANCEMENT
; NUMBER OF SEQUENCES: 33
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,305
; FILING DATE: 18-NOV-1996
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; COMPUTER TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,305
; FILING DATE: 18-NOV-1996
; ADDRESSEE: Garcia-Rodriguez, Consuelo
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,021
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,007
; FILING DATE: 30-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/848,261
; FILING DATE: 29-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett, Arthur S.
; REGISTRATION NUMBER: 20,338
; REFERENCE/DOCKET NUMBER: 06899.0001-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 9;
; SEQUENCE CHARACTERISTICS:

RESULT 14
US-08-627-706-15
; Sequence 15, Application US/08627706
; Patent No. 5773696

GENERAL INFORMATION:
 APPLICANT: Liang, Jihong
 APPLICANT: Shah, Dilip M.
 APPLICANT: Wu, Yonnie S.
 APPLICANT: Rosenberger, Cindy A.
 TITLE OF INVENTION: Antifungal Polypeptide and Methods for Controlling Plant Pathogenic Fungi
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Charles E. Cohen, Monsanto Company, BB4F
 STREET: 700 Chesterfield Village Parkway No. 577369th
 CITY: St. Louis
 STATE: Missouri
 COUNTRY: USA
 ZIP: 63198
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 REFERENCE/DOCKET NUMBER: US/09/103,489
 FILING DATE: 24-JUN-1998
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Cohen, Charles E.
 REGISTRATION NUMBER: 34,565
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (314) 537-6224
 TELEFAX: (314) 537-6047
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 79 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 CLASSIFICATION: 436
 ATTORNEY/AGENT INFORMATION:
 NAME: Cohen, Charles E.
 REGISTRATION NUMBER: 34,565
 REFERENCE/DOCKET NUMBER: 38-21(10700)A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (314) 537-6224
 TELEFAX: (314) 537-6047
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 79 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: Peptide
 * US-09-627-706-15

Query Match 17.7%; Score 60; DB 4;
 Best Local Similarity 27.5%; Pred. No. 2,7; Mismatches 9; Indels 12; Gaps 4;
 Matches 19; Conservative 9;
 QY 4 IAIIFIVLYAFCLLE-DGIVEAGF-----GCPENAGKCHRHKSIIRR-RGGFCRGTR 52
 Db 8 ISLIFAAALVLFAAFEAPTMDARLCERPSGTWSGVCGNNNACRNQCRNLERAEHGSCNYV 67
 QY 53 FRT-TCVCY 60
 Db 68 FPAHKCICY 76

Search completed: September 17, 2002, 15:44:35
 Job time: 126 sec

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 REFERENCE/DOCKET NUMBER: US/09/103,489
 FILING DATE: 24-JUN-1998
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Cohen, Charles E.
 REGISTRATION NUMBER: 34,565
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (314) 537-6224
 TELEFAX: (314) 537-6047
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 79 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-09-103-489-15

Query Match 17.7%; Score 60; DB 4;
 Best Local Similarity 27.5%; Pred. No. 2,7; Mismatches 9; Indels 12; Gaps 4;
 Matches 19; Conservative 9;
 QY 4 IAIIFIVLYAFCLLE-DGIVEAGF-----GCPENAGKCHRHKSIIRR-RGGFCRGTR 52
 Db 8 ISLIFAAALVLFAAFEAPTMDARLCERPSGTWSGVCGNNNACRNQCRNLERAEHGSCNYV 67
 QY 53 FRT-TCVCY 60
 Db 68 FPAHKCICY 76

RESULT 15
 US-09-103-489-15
 ; Sequence 15, Application US/09/03489
 ; Patent No. 6215048
 ; GENERAL INFORMATION:
 ; APPLICANT: Liang, Jihong
 ; APPLICANT: Shah, Dilip M.
 ; APPLICANT: Wu, Yonnie S.
 ; APPLICANT: Rosenberger, Cindy A.
 ; TITLE OF INVENTION: Antifungal Polypeptide and Methods for Controlling Plant Pathogenic Fungi
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Charles E. Cohen, Monsanto Company, BB4F
 ; STREET: 700 Chesterfield Village Parkway No. 6215048th
 ; CITY: St. Louis
 ; STATE: Missouri
 ; COUNTRY: USA
 ; ZIP: 63198

Result No.	Score	Query Match	Length	DB ID	Description	%
1	177	52.2	38	19 AAW66437	Cationic Peptide s	
2	177	52.2	38	21 AA91736	Plant SDF encoded	
3	133.5	39.4	38	15 AAR51160	Plant SDF encoded	
4	133.5	39.4	38	19 AAW66436	Antibacterial poly	
5	133.5	39.4	38	21 AA91735	Plant SDF encoded	
6	71.5	21.1	43	16 AAR82841	Teneocin. Teneocin	
7	70	20.6	75	22 AAE10361	Plant SDF encoded	
8	70	20.6	94	10 AAF94261	Plant SDF encoded	
9	69.5	20.5	43	16 AAR82842	Antibacterial poly	
10	68.5	20.2	40	14 AAR30562	Allomycin B antiba	
11	68.5	20.2	72	21 AAY44509	Sapecin - an antib	
					Pea Defensin prote	
					New indolicidin peptide analogues - useful for, e.g. enhancing	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:42:09 ; Search time 29.91 Seconds
(without alignments)
226,530 Million cell updates/sec

Title: US-09-829-481-4
Perfect score: 339
Sequence: 1 MKSIAIIIFTWVAFCAILEDGG.....IRRRGGFCRGIFRTTCYCYR 61

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802:
 1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1980.DAT: *
 2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1981.DAT: *
 3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1982.DAT: *
 4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1983.DAT: *
 5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1984.DAT: *
 6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1985.DAT: *
 7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1986.DAT: *
 8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1987.DAT: *
 9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1988.DAT: *
 10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1989.DAT: *
 11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1990.DAT: *
 12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1991.DAT: *
 13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1992.DAT: *
 14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1993.DAT: *
 15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1994.DAT: *
 16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1995.DAT: *
 17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1996.DAT: *
 18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1997.DAT: *
 19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1998.DAT: *
 20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA1999.DAT: *
 21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA2000.DAT: *
 22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb/AA2001.DAT: *

ALIGNMENTS

RESULT 1
AAW66437
ID AAW66437 standard; peptide; 38 AA.
XX
AC AAW66437;
XX
XX 12-JAN-1999 (first entry)

cationic peptide scorpion defensin.

XX DE
XX
XX KW Indolicidin analogue; resistance; cationic peptide; antibiotic; bacterial infection; tolerance; antibiotic; microorganism; KW bacteria; fungus; parasite; virus.
XX OS Deirus quinquestriatus.
XX PN WO9840401-A2.

17-SEP-1998.

PD

XX

PP

XX

PR

XX

PR

XX

MICR- MICROLOGIX BIOTECH INC.

Fraser JR, McNicol PJ, West MHP;

WPI; 1998-520800/44.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	%
1	177	52.2	38	19 AAW66437	Cationic Peptide s	
2	133.5	39.4	38	15 AAR51160	Cationic Peptide s	
3	133.5	39.4	38	19 AAW66436	Antibacterial pept	
4	133.5	39.4	38	21 AA91735	Cationic Peptide i	
5	133.5	39.4	38	16 AAR82841	Cationic Peptide i	
6	71.5	21.1	43	16 AAR82841	Allomycin A antiba	
7	70	20.6	75	22 AAE10361	Soybean Gly m2 pro	
8	70	20.6	94	10 AAF94261	Antibacterial poly	
9	69.5	20.5	43	16 AAR82842	Allomycin B antiba	
10	68.5	20.2	40	14 AAR30562	Sapecin - an antib	
11	68.5	20.2	72	21 AAY44509	Pea Defensin prote	

PT activity of antibiotic or overcoming tolerance, acquired resistance
 PT or inherent resistance of microorganisms
 XX Disclosure; Page 10; 105pp; English.
 XX AAW66469 to AAW66469 represent native cationic peptides from the present invention. The present invention describes compositions and methods for treating infection, especially bacterial infections. The compositions and methods use cationic peptides in combination with an antibiotic agent which are then administered to a patient to enhance the activity of the antibiotic agent, to overcome: (a) tolerance; (b) acquired resistance; and (c) inherent resistance. The combinations of antibiotics and cationic peptides can provide synergistic activity against a microorganism that is tolerant, inherently resistant, or has acquired resistance to an antibiotic agent. They can be used for killing e.g. bacteria, fungi, parasites and viruses.

XX Sequence 38 AA;

Query Match 52.2%; Score 177; DB 19; Length 38;
 Best Local Similarity 73.0%; Pred. No. 1.9e-13;
 Matches 27; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 25 GFGCPNAGKCHRHKSKIRRGFCRGTFRITCVYR 61
 Db 1 gfgcpnqgachrcsrrrgycagffkqtctvyr 37

RESULT 2
 AAY91736 standard; Peptide: 38 AA.
 XX AAY91736;
 XX 06-JUN-2000 (first entry)
 DE Cationic peptide scorpion defensin amino acid sequence.
 XX Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukaemia; polyoxalkylene-modified APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KW multidrug resistance.
 XX Unidentified.
 XX WO9965506-A2.
 XX PD 23-DEC-1999.
 XX PF 14-JUN-1999; 99WO-CA00552.
 XX PR 12-JUN-1998; 98US-0096511.
 XX (MICR-) MICROLOGIX BIOTECH INC.

PT Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;

XX DR WPI: 2000-223549/19.

XX Novel pharmaceutical composition containing optionally activated polyoxalkylene-modified cationic peptides, useful for treating tumours

XX Disclosure; Page 11; 94pp; English.

XX This sequence represents a cationic peptide amino acid sequence, which can be used in the pharmaceutical composition of the invention. The invention relates to a pharmaceutical composition containing at least one activated polyoxalkylene (APO)-modified cationic peptide. The

CC modification of peptides with APO increases their activity against tumour

CC cells, including those with a multidrug resistant phenotype. The

CC pharmaceutical composition can be used to treat tumours, specifically

CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary, cervix, uterus, skin, prostate, liver and colon.
 XX SQ Sequence 38 AA;
 XX Query Match 52.2%; Score 177; DB 21; Length 38;
 CC Best Local Similarity 73.0%; Pred. No. 1.9e-13;
 CC Matches 27; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 CC Qy 25 GFGCPNAGKCHRHKSKIRRGFCRGTFRITCVYR 61
 CC Db 1 gfgcpnqgachrcsrrrgycagffkqtctvyr 37

RESULT 3
 AAR51160
 ID AAR51160 standard; Peptide: 38 AA.
 XX AC AAR51160;
 XX AC AAR51160;
 XX DT 18-OCT-1994 (first entry)
 XX DE Antibacterial peptide induced in Aeschna cyanea.
 XX KW Defensin; antibacterial peptide; Odonata; paleopteran insect;
 XX KW dragonfly.
 OS Aeschna cyanea.
 XX FH Key Disulfide-bond 4..26
 FT Disulfide-bond 11..34
 FT Disulfide-bond 14..36
 FT Region 11..21
 FT /label= "alpha_helix
 FT /note= "approximate position"
 FT Region 24..38
 FT /label= "beta-pelated_sheet
 FT /note= "approximate position"
 XX PN FR2695392-A.
 PD 11-MAR-1994
 XX PF 04-SEP-1992; 92FR-0010609.
 XX PR 04-SEP-1992; 92FR-0010609.
 XX PA (CNRS) CNRS CENT NAT RECH SCI.
 XX PI Bulet P, Hetru C, Hoffmann J;
 XX DR WPI: 1994-103535/13.
 XX PT New antibacterial peptide(s) from dragonfly - for medical,
 PT veterinary, agricultural and food preservation use
 XX PS Claim 10; Page 17; 25pp; French.
 XX CC A defensin is induced in the dragonfly Aeschna cyanea by bacterial
 CC CC infection. The defensin (AAR51160) contains 3 disulphide bonds and has
 CC a structure comprising an alpha helix linked to a C-terminal
 CC antiparallel beta-peleted sheet through 2 of the disulphide bonds.
 CC The peptide has antibacterial activity which makes it suitable for
 CC use in the treatment of bacterial infections of eyes and ears and
 CC for oral/dental hygiene and gynaecology.
 XX SQ Sequence 38 AA;

Query Match 39.4%; Score 133.5; DB 15; Length 38;
 Best Local Similarity 55.3%; Pred. No. 1.e-08;

XX	PF	28-DEC-1993;	93JP-0350294.		PT allergen, useful for producing soybean plants which can be used to make soybean products -	
XX	PR	28-DEC-1993;	93JP-0350294.		XX Claim 56; Page 52; 57pp; English.	
PA	(AMAN) AMANO PHARM KK.				XX	
XX	DR	1995-299564/39.			CC The patent discloses hypoallergenic transgenic soybeans and recombinant expression constructs to lower soybean vacuolar protein, commonly known as P34, (Gly m Bd 30K or Gly m 1) and other allergens such as Gly m IA, Gly m IB, rGLY m3 or Glycinin G1 (alabib). The allergen content of the soybean is reduced by sense suppression which is accomplished by using the expression construct that comprises a nucleic acid fragment encoding the allergen. The constructs are useful for producing hypoallergenic transgenic soybean plants which can be used to make hypoallergenic soybean products which can be used in a variety of food (e.g., infant formulas) and animal feed applications. The oil made from seeds of the soybean plants can be used as ingredients, as coatings, as salad oils, as spraying oils, and as frying oils. The foods in which the oil may be used include, crackers and snack foods, confectionery products, syrups and toppings, sauces, batter and breading mixtures, baking mixes and doughs. The present sequence is soybean Gly m2 protein from s1s1c_pk027_all clone. This sequence is a minor soybean seed allergen.	
DR	N-PSDB; AAT01306.				CC	
XX	Novel polypeptide derived from Allomyrina dichotoma larvae - is physiologically active against bacteria, and may be used in drugs and foodstuffs				CC	
XX	PS	Claim 1; Page 10; 14pp; Japanese.			CC	
XX	Two new peptides have been derived from Allomyrina dichotoma, allomycin A and allomycin B (AAR81841 and AAR82842). The new peptides have antibacterial activity which makes them useful in the drug and foodstuff industries. The proteins may be prepared using DNA coding for the peptides (AAT01306 and AAT01307) in standard recombinant techniques, thus also enables the large scale production of large quantities of the peptides.				CC	
XX	Sequence	43 AA;			CC	
XX					CC	
Query Match	Score 21.1%;	Score 71.5%;	Score 70.0%;	Score 20.6%;	Query Match	
Best Local Similarity	42.9%;	DB 16;	DB 22;	DB 22;	Best Local Similarity	
Matches	Pred. No. 0.28;	Length 43;	Length 75;	Length 75;	Matches	
Matches 15; Conservative	2; Mismatches 15;	Indels 3;	Gaps 1;	Indels 18;	Best Local Similarity 32.4%; Pred. No. 0.73; Gaps 4;	
Qy	27 GCPFNAGKCHHCKSIRRGGFCRTFRITCVCYR 61				Qy 2 KSIA-LIFIVLVAFCLLE-----DGIVEAGFGCPFMAGKCHRCKS---IRRGG 46	
Db	12 9faahns1cahclraigrrgcs---rgvcicrr 43				Db 4 ksiagifclflvfaqevvvqteaktcenladtyrgpcfttgsodchcknkehllr--- 60	
RESULT	7				Qy 47 GFCRGTFTRITC 57	
ID	AAE10361	standard; Protein; 75 AA.			Db 61 grcddfrcw 71	
XX	AAE10361;					
XX	DT	10-DEC-2001	(first entry)		RESULT 8	
XX	DE	Soybean Gly m2 protein from clone s1s1c_pk027_all.			AAP94261	
XX	Soybean; allergen; transgenic plant; P34 protein; Gly m Bd 30K; Gly m 1; soybean vacuolar protein; Gly m IA; rGLY m3; Glycinin G1; alabib; food; infant formula; coating; salad oil; syrup; spraying oil; roasting oil; frying oil; cracker; confectionery product; snack food; topping; sauce; batter; breading mixture; baking mix; dough; Gly m2 protein; clone s1s1c_pk027_all.			ID	AAP94261 standard; protein; 94 AA.	
XX	OS	Glycine max.			XX	
XX	PN	WO200568887-A2.			AC	AAP94261;
XX	PD	20-SEP-2001.			DT	17-JUN-1990 (first entry)
XX	PR	15-MAR-2001; 2001WO-US082254.			XX	Antibacterial polypeptide precursor of Sarcophaga peregrina.
XX	PR	16-MAR-2000; 2000US-0189823.			DE	Antibacterial polypeptide precursor of Sarcophaga peregrina.
XX	PA	(DUPO) DU PONT DE Nemours & CO E I .			XX	KW Antibacterial polypeptide precursor; Sarcophaga peregrina.
XX	PA	(PION) PIONEER HI-BRED INT INC.			OS	XX Sarcophaga peregrina.
XX	PI	Jung R, Kinney AJ;			FH	XX Key Location/Qualifiers
XX	PR	WPI; 2001-582460/65.			FT	55..94 FT /note="cloned antibacterial polypeptide"
XX	PR	N-PSDB; AAD17530.			FT	24..52 FT /note="previously disclosed sequence"
XX	PR	Recombinant expression construct to lower allergen (e.g. Gly m Bd 30K) content of a soybean, comprises a nucleic acid fragment encoding the			FT	1..23 FT /label=signal sequence
XX	PA	(SANWA KAGAKU KENKYUSHO CO. LTD.			XX	XX EP303858-A.
XX	PD	22-FEB-1989.			XX	XX 22-FEB-1989,
XX	PR	26-JUL-1988;			XX	XX 26-JUL-1988;
XX	PR	20-AUG-1987;			XX	XX 20-AUG-1987;
XX	PA	(SANWA) SANWA KAGAKU KENKYUSHO CO. LTD.			XX	XX PA
XX	PI	Sawai K, Natori S, Takahashi H, Tanaka K, Mitanai T, Kurono M;			XX	XX PI
XX	PR				XX	XX PR

PS Disclosure; Page 11; 94pp; English.
 XX This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalkylene (APO)-modified cationic peptide. The
 CC modification of Peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.
 XX Sequence 40 AA;

Query Match 20.1%; Score 68; DB 21; Length 40;
 Best Local Similarity 39.4%; Pred. No. 0.65; Gaps 1;
 Matches 13; Conservative 2; Mismatches 16; Indels 2; Gaps 1;

Qy 27 GCPFNAGKCHRHCKSIRRGFCRATFRITCVC 59
 Db 8 gtginhsacaaahcllrrgnrgcng--kavcvc 38

RESULT 14
 ID AAB24719 standard; Peptide: 77 AA.
 XX AC AAB24719;
 XX DT 27 -NOV-2000 (first entry)
 XX DE Plant SDF encoded polypeptide sequence SEQ List 1 NO:162.
 XX KW Plant; corn; Arabidopsis thaliana; sequence-determined DNA fragment;
 KW SDF; genetic mapping; identification; promoter; structural gene; UTR;
 KW untranslated region; expression control.
 XX OS Plant.
 XX PN WO2000040695-A2.
 XX PR 08 -JAN-1999; 99US-0115293.
 XX PA (CERE-) CERES INC.
 XX PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhman ME;
 XX DR 07 -JAN-2000; 2000WO-US00466.
 XX PR 08 -JAN-1999; 99US-0115293.
 XX PA (CERE-) CERES INC.
 XX PI Zheng L;
 XX DR 2000-465970/40.
 XX PT New corn plant and Arabidopsis thaliana sequence-determined DNA
 PT fragments, useful for expressing gene products and for controlling
 PT expression of a target gene -
 XX PS Claim 14: Page 400; 673pp; English.

XX The present invention describes polynucleotides, such as complete cDNA
 CC sequences and/or sequences of genomic DNA encompassing complete genes,
 CC portions of genes, and/or intergenic regions, collectively referred to
 CC as sequence-determined DNA fragments (SDFs), from corn plants and
 CC Arabidopsis thaliana. The SDFs are promoters, structural genes,
 CC untranslated regions (UTRs), or 3' termination sequences. They can be
 CC used for expressing a gene product and controlling expression of a
 CC target gene, either as a promoter, a structural gene, an UTR or as a
 CC 3' termination sequence. They are also useful as tools for genetic
 CC mapping, and identification of a particular individual plant or for
 CC clustering a group of plants with a common trait. AAA78433 to AAA78630
 CC and AA24605 to AA25099 represent the specifically claimed
 CC polynucleotide sequences and polypeptides encoded by them given in the
 CC present invention.
 XX SQ Sequence 78 AA;

CC polynucleotide sequences and polypeptides encoded by them given in the
 CC present invention.
 XX SQ Sequence 77 AA;

Query Match 20.1%; Score 68; DB 21; Length 77;
 Best Local Similarity 30.1%; Pred. No. 1.3;
 Matches 22; Conservative 9; Mismatches 26; Indels 16; Gaps 4;
 Qy 2 KSIAlIIFIVLVAFCLIEDGV-EAGFGCPEN-----AGKCHRHKCSIRRG-G 47
 Db 3 klgqvfrwmiiitlvlgvvaneglqpkkgneilqgsncvaecdsmcvkkrgkag 62
 Qy 48 FCRGTFRTICVCY 60
 Db 63 ycspskk -cycy 73

RESULT 15
 ID AAB24718 standard; Peptide: 78 AA.
 XX AC AAB24718;
 XX DT 27 -NOV-2000 (first entry)
 XX DE Plant SDF encoded polypeptide sequence SEQ List 1 NO:161.
 XX KW Plant; corn; Arabidopsis thaliana; sequence-determined DNA fragment;
 KW SDF; genetic mapping; identification; promoter; structural gene; UTR;
 KW untranslated region; expression control.
 XX OS Plant.
 XX PN WO2000040695-A2.
 XX PR 08 -JAN-1999; 99US-0115293.
 XX PA (CERE-) CERES INC.
 XX PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhman ME;
 XX DR 07 -JAN-2000; 2000WO-US00466.
 XX PR 08 -JAN-1999; 99US-0115293.
 XX PA (CERE-) CERES INC.
 XX PI Zheng L;
 XX DR 2000-465970/40.
 XX PS Claim 14: Page 399; 673pp; English.
 XX The present invention describes polynucleotides, such as complete cDNA
 CC sequences and/or sequences of genomic DNA encompassing complete genes,
 CC portions of genes, and/or intergenic regions, collectively referred to
 CC as sequence-determined DNA fragments (SDFs), from corn plants and
 CC Arabidopsis thaliana. The SDFs are promoters, structural genes,
 CC untranslated regions (UTRs), or 3' termination sequences. They can be
 CC used for expressing a gene product and controlling expression of a
 CC target gene, either as a promoter, a structural gene, an UTR or as a
 CC 3' termination sequence. They are also useful as tools for genetic
 CC mapping, and identification of a particular individual plant or for
 CC clustering a group of plants with a common trait. AAA78433 to AAA78630
 CC and AA24605 to AA25099 represent the specifically claimed
 CC polynucleotide sequences and polypeptides encoded by them given in the
 CC present invention.
 XX SQ Sequence 78 AA;

Query	Match	20.1%	Score	68;	DB	21;	Length	78;
Best Local Similarity	30.1%	No.	1.3;	Pred.				
Matches	22;	Conservative	9;	Mismatches	26;	Indels	16;	Gaps
QY	2	KSLIALLIVAFACILEDGLV-EAGFGCPFN-----						
Ddb	4	kliqvsftwmiaiftlvgvaneglgkpkkqcneillqsnhevaaecdsmcvkkrqgaa						
QY	48	FCRGTTERTTCVCY	60					
Ddb	64	ycspsskk--ccyccy	74					

Search completed: September 17, 2002, 15:44:15
Job time: 126 sec

GenCore version 4.5			
Copyright (c) 1993 - 2000 CompuGen Ltd.			
protein - protein search, using sw model			
on:	September 17, 2002, 15:43:44 ;	Search time 10.33 Seconds (without alignments)	
file:	US-09-829-481-4	228.644 Million cell updates/sec	
Sequence:	1 MKSTAIIFIVLVAFCILEDG	IRRGGFCRGTFRTVCYR	61
Scoring table:	BLOSUM62	GapOp 10.0 , GapExt 0.5	
Searched:	105224 seqs,	38719550 residues	
Total number of hits satisfying chosen parameters:	105224		
Maximum DB seq length: 0			
Minimum DB seq length: 2000000000			
st-processing: Minimum Match 0%			
st-processing: Maximum Match 100%			
Listing first 45 summaries			
Database :	SwissProt_40:*		
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.			
SUMMARIES			
result No.	Score	Query Length	DB ID
1	184	54.3	DEF4_ANDAU
2	177	52.2	DEF4_ANDAU
3	133.5	39.4	DEF1_AECSI
4	128	37.8	DEF1_MYTGA
5	127	37.5	DEF1_MYTGD
6	114	33.6	DEFB_MYTED
7	82.5	24.3	SC11_MEWSA
8	80.5	23.7	DEF1_ZOPAT
9	79.5	23.5	SCX2_ANDAU
10	75.5	22.3	SC11_MEWSA
11	71.5	21.1	DEF1_ALLDI
12	71.5	21.1	DEF2_CAPAN
13	71.5	21.1	DEF1_TENMO
14	71	9.9	DEF1_AEDAE
15	70.5	20.8	THGF_HELAN
16	69.5	20.5	SCXB_BUTOC
17	69	20.4	1OKD_VIGUN
18	69	20.4	DEF1_ANOGA
19	68.5	20.2	D230_PEA
20	68	20.2	SC15_MEWSA
21	68	20.1	SAPE_SARPE
22	66.5	19.6	SCX3_BUTOC
23	66.5	19.6	SCX7_CENNO
24	66	19.5	DEF1_PROTE
25	65.5	19.3	SCX5_ANDMA
26	65.5	19.3	SCX5_ANNDMA
27	64.5	19.3	SCX5_LIEQH
28	63.5	18.7	ITB7_HUMAN
29	63.5	18.7	ITB7_ARATH
30	63.5	18.7	ITB7_RAPSA
31	63.5	18.7	SCRP_PANIM
32	63.5	18.7	TITYU_BAHII
33	63	18.6	ITB8_RABBIT
34	63	18.6	ITI4_ARATH
35	61.5	18.1	ITG1_ARATH
36	61.5	18.1	APP1_BRANA
37	61.5	18.1	SCX1_CENNO
38	61	18.0	SAFC_SARPE
39	60.5	17.8	APP1_ARATH
40	60.5	17.8	Q9RU01_deinococcus
41	60.5	17.8	P26011_mus_musculus
42	60	17.7	ITB7_MOUSE
43	60	17.7	APP3_RAPSA
44	60	17.7	P56611_tityu_bahi
45	60	17.7	Q9UAC8_mesobuthus
			P09617_hordeum_vul
ALIGNMENTS			
RESULT 1			
DEF4_ANDAU	STANDARD;	PRT;	37 AA.
ID DEF4_ANDAU	P81618;		
AC P86686;			
DT 15-JUL-1999	(Rel. 38, Created)		
DT 15-JUL-1999	(Rel. 38, Last sequence update)		
DT 30-MAY-2000	(Rel. 39, Last annotation update)		
DE 4 kDa defensin.			
OS Androctonus australis hector (Sahara scorpion).			
OC Metazoa; Arthropoda; Chelicera; Arachnida; Scorpiones; Buthoidea; Buthidae; Androctonus.			
NCBI_TAXID=70175;			
RN [1]			
SEQUENCE, AND CHARACTERIZATION.			
RP TISSUE-Hemolymph;			
RC MEDLINE#97054646; PubMed=8939880;			
RA Ehret-Sabatier L., Low D., Goyffon M., Fehlbaum P., Hoffmann J.A., van Dorselaer A., Bulet P.			
RT "Characterization of novel cysteine-rich antimicrobial peptides from scorpion blood";			
RL J. Biol. Chem. 271:29537-29544(1996).			
CC [-] FUNCTION: ACTIVE AGAINST GRAM-POSITIVE BACTERIA.			
CC [-] SUBCELLULAR LOCATION: Secreted.			
CC [-] MASS SPECTROMETRY: MW=4206.8; METHOD=Electrospray.			
CC [-] SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.			
DR HSSP; P10891; ICAC.			
DR InterPro; IPR001542; Arthro_defensin.			
DR PF01037; Arthro_defensin.			
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.			
KW Antibiotic.			
FT DISULFID	4	25	BY SIMILARITY.
FT DISULFID	11	33	BY SIMILARITY.
FT DISULFID	15	35	BY SIMILARITY.
SQ SEQUENCE	37 AA;	4212 MW;	AB13638EGE3FB84C1 CRC64;
Query Match	54 / 38;	Score 184;	DB 1;
Best Local Similarity	75.7 %;	Pred. NO. 2.2-e-1;	
Matches	28;	Conservative	3; Mismatches 6;
		Indels	0; Gaps
RESULT 2			
DEF4_LIEQH	STANDARD;	PRT;	38 AA.
ID DEF4_LIEQH			
AC P41965;			
DT 01-NOV-1995	(Rel. 32, Created)		
DT 01-NOV-1995	(Rel. 32, Last sequence update)		
DT 01-NOV-1995	(Rel. 32, Last annotation update)		
DE 4 kDa defensin (Antibacterial 4 kDa peptide).			
OS Leurus quinquestrigatus hebraeus (Scorpion); Buthoidea; Buthidae; Leurus.			
OC Mesobuthus; Buthidae; Leurus.			
OC Buthoidea; Buthidae; Leurus.			

RL	J. Biol. Chem. 271:21808-21813(1996).			
- - FUNCTION:	HAS ANTIBACTERIAL ACTIVITY AGAINST M. LUTEUS AND E. COLI.			
CC	- - SUBCELLULAR LOCATION: Secreted.			
CC	- - MASS SPECTROMETRY: MW=4392.; METHOD=MALDI.			
CC	- - SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.			
CC	DR InterPro: IPR001542; Arthro_defensin.			
DR	Pfam: PF01097; Arthro_defensin; 1.			
DR	PROSITE: PS00425; ARTHROPOD_DEFENSINS; 1.			
KW	Antibiotic.	4	25	BY SIMILARITY.
FT	DISULFID	10	33	BY SIMILARITY.
FT	DISULFID	14	35	BY SIMILARITY.
SO	SEQUENCE	37 AA;	4149 MW;	8DEDCCFOE78E6DA1 CRC64;
Qy	25 GFGCPFNAGGCHRCKST-RERRGFCRGTFRTCVCYR 61			
Db	1 GFGCP-NDYCHRHICKSIPQRXGSGXHRLRCTCYR 37			
RESULT	6			
ID	DEFB_MYTED	STANDARD;	PRT;	35 AA.
AC	P81611;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DE	DEFENSIN B (Fragment).			
OS	Mytilus edulis (Blue mussel).			
OC	Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Mytiloidea;			
CC	Mytiloidea; Mytilidae; Mytilus.			
OX	NCBI_TAXID=6550;			
RN	[1]			
RP	SEQUENCE, AND CHARACTERIZATION.			
RC	TISSUE-Blood;			
RX	MedlineP=9635569; PubMed=8702979;			
RA	Charlet M., Chernysh S., Philippe H., Hetru C., Hoffman J. A., Bulet P.;			
RA	"Innate immunity. Isolation of several cysteine-rich antimicrobial peptides from the blood of a mollusc, <i>Mytilus edulis</i> ."			
RT	J. Biol. Chem. 271:21808-21813(1996).			
CC	- - FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST M. LUTEUS AND E. COLI.			
CC	- - SUBCELLULAR LOCATION: Secreted.			
CC	- - MASS SPECTROMETRY: MW=4314.3.; METHOD=MALDI.			
CC	DR InterPro: IPR001542; Arthro_defensin.			
DR	Pfam: PF01097; Arthro_defensin; 1.			
DR	PROSITE: PS00425; ARTHROPOD_DEFENSINS; 1.			
KW	Antibiotic.	4	25	BY SIMILARITY.
FT	DISULFID	10	33	BY SIMILARITY.
FT	DISULFID	14	35	BY SIMILARITY.
FT	NON_TER	35	35	BY SIMILARITY.
SO	SEQUENCE	35 AA;	3882 MW;	5CF0E78E6DA1B8D3 CRC64;
Query Match	33.68%; Score 114; DB 1; Length 35;			
Best Local Similarity 61.1%; Pred. No. 1..4e-06;				
Matches 22; Conservative 1; Mismatches 11; Indels 2; Gaps 2;				
Qy	25 GFGCPFNAGKCHRHCKST-RRRGFCRFRTCVCYR 59			
Db	1 GFGCP-NDYCHRHCKSIPGRGGCGXHRLRCTC 35			
RESULT	7			
SC11_MEWSA	STANDARD;	PRT;	43 AA.	

ID SC11_MEWSA STANDARD; PRT; 85 AA.

AC Q9NJJC7; Q9NJJP8; ID DEFA_ZOPAT

DT 01-MAR-2002 (Rel. 41, Created) ID DEFA_ZOPAT

DT 01-MAR-2002 (Rel. 41, Last sequence update) ID DEFA_ZOPAT

DT 01-MAR-2002 (Rel. 41, Last annotation update) ID DEFA_ZOPAT

DE Alpha neurotoxin TX11' (Alpha toxin 2). ID DEFA_ZOPAT

OS Mesobuthus martensi (Manchurian scorpion) (Buthus martensi). ID DEFA_ZOPAT

OC Eukaryota; Metazoa; Arachnida; Arachicerata; Cheliceraida; Scorpiones; ID DEFA_ZOPAT

OC Buthoidea; Butidae; Buthus; ID DEFA_ZOPAT

OX NCBI_TAXID=34649; OX NCBI_TAXID=34649;

RN [1] RQ SEQUENCE FROM N. A.

RC TISSUE-Venom gland; MEDILINEP=20317244; PubMed=10858508; RX Zhu S.-Y., Li W.-X., Zeng X.-C., Liu H., Jiang D.-H., Mao X.; RT "Nine novel precursors of *Buthus martensi* scorpion alpha-toxin homologues." RT Toxicon 38:1653-1661(2000). RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

CC SEQUENCE FROM N. A.

RA Ye J., Chen J., Zuo X., Ji Y.; "Cloning and characterization of cDNA sequences encoding two novel alpha toxin precursors from the Chinese scorpion *Buthus martensis* Karsch"; RQ Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/ or send an email to license@isb-sib.ch).

CC DR EMBL; AF15534; AAF31418; 1; DR EMBL; AF132974; AAF31295; 1; DR EMBL; AF288608; AAG00581; 1; DR HSSP; AF01484; 1AHO; DR InterPro; IPR003614; Knot1; DR InterPro; IPR00206; Scorpion_toxin. DR InterPro; IPR001219; Toxin. DR PRINTS; PRO0284; TOXIN. DR ProDom; PD000308; Scorpion_toxin; 1. DR SMART; SM00505; Knot1; 1. KW Neurotoxin; Sodium channel inhibitor; Signal. KW SIGNAL 1 19 POTENTIAL. FT CHAIN 20 85 ALPHA-NEUROTOXIN TX11.

FT DISULFID 31 82 BY SIMILARITY.

FT DISULFID 35 55 BY SIMILARITY.

FT DISULFID 41 65 BY SIMILARITY.

FT DISULFID 45 67 BY SIMILARITY.

FT VARIANT 4 4 VVII->LVPE (IN ISOFORM TX11P).

FT VARIANT 13 13 V->L (IN ISOFORM TX11P).

SQ SEQUENCE 85 AA; 9524 MW; 565390BB6E71806E CRC64;

Query Match Score 24.3%; DB 1; Length 85; Best Local Similarity 25.8%; Pred. No. 0.0088; Matches 17; Mismatches 13; Indels 5; Gaps 2;

Qy 1 MKSIAIIFIVLVAFCILEDGIVEAGFCGPENAGK --CHRHKCKSTRRRGGFCR -GTFRT 55

Db 4 MVIISIALLYMTGVBSVKDGYIADDRNCPYFCGRNAYCDGECKKNRAESGYCOMASKYGN 63

RESULT 8

DEFA_ZOPAT

ID DEFA_ZOPAT

AC P80033;

DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 01-MAR-1992 (Rel. 21, Last annotation update)
 DE Defensin; isoforms B and C.
 OS Zophobas atratus.
 OC Pterigota: Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Coleoptera; Tenebrionidae; Zophobas.
 RN [1] NCBI_TAXID=7074;
 RP SEQUENCE.

RC TISSUE="Hemolymph";
 RX MEDLINE=92105112; PubMed=1761552;
 RA Bulet P., Cocianich S., Dinarcq J.-L., Lambert J., Reichhart J.-M., Hoffmann D., Hetru C., Hoffmann J.A., Hoffmann D.;
 RT "Insect immunity. Isolation from a coleopteran insect of a novel inducible antibacterial peptide and of new members of the insect defensin family."
 RT J. Biol. Chem., 266:24520-24525 (1991).
 CC FUNCTION: INVOLVED IN ANTI GRAM-POSITIVE ACTIVITY OF IMMUNE DEFOMYMPH OF Z. ATRATUS.
 CC 1- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.
 DR PIR: B41711; BA1711.
 DR PIR: C41711; C41711.
 DR HSSP: P10891; IICAA.
 DR InterPro: IPR01542; Arthro_defensin.
 DR InterPro: IPR03614; Knot1.
 DR Pfam: PF01097; Arthro_defensin; 1.
 DR PRINTS; PRO0271; DEFENSINS; 1.
 DR SMART; SM00505; Knot1; 1.
 DR PROSITE; PS000425; ARTHROPOD_DEFENSINS; 1.
 KW Insect immunity; Antibiotic.
 FT DISULFID 3 34 BY SIMILARITY.
 FT DISULFID 20 40 BY SIMILARITY.
 FT DISULFID 24 42 BY SIMILARITY.
 FT VARIANT 30 30 R->T (IN ISOFORM C).
 SQ SEQUENCE 43 AA; 4453 MW; D63E7B100987FD2F CRC64;

Query Match 9
 Best Local Similarity Score 80.5; DB 1; Length 43;
 Matches 15; Conservative 6; Mismatches 12; Indels 5; Gaps 2;

Qy 25 GF-->GCPFNAGKCHRHKCSIRRGGFCRGTFERTTCYC 59
 Db 7 GFEIAGTKNSAANGAHLAQLGRGGCNS - KSVCVVC 42

RESULT 9
 ID SCX2_ANDAU STANDARD; PRT; 85 AA.
 AC P0148a;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurotoxin II Precursor (Aah II).
 OS Androctonus australis Hector (Sahara scorpion).
 OC Eukaryota: Metazoa; Arthropoda; Chelicerata; Scorpiones; Buthoidea; Butidae; Androctonus.
 OX NCBI_TAXID=70175;
 RN [1] SEQUENCE FROM N.A.
 RP MEDLINE=90037002; PubMed=2808423;
 RA Bougis P.E., Rochat H., Smith L.A.;
 RT "Precursors of Androctonus australis scorpion neurotoxins. Structures of precursors, processing outcomes, and expression of a functional recombinant toxin II.";
 RL J. Biol. Chem. 264:19259-19265 (1989).
 RN [2] SEQUENCE OF 20-82.
 RP MEDLINE=73025153; PubMed=4312910;
 RA Rochat H., Rochat C., Sampieri F., Miranda F., Lissitzky S.;
 RT "The amino-acid sequence of neurotoxin II of Androctonus australis

hector.";
 Eur. J. Biochem. 28:381-388 (1972).
 RT RL
 RN [3] DISULFIDE BONDS.
 RP MEDLINE=75057756; PubMed=4611766;
 RX RA Kopayan C., Martinez G., Lissitzky S., Miranda F., Rochat H.;
 RT RT "Disulfide bonds of toxin II of the scorpion Androctonus australis hector.";
 RP RL
 RL Eur. J. Biochem. 47:488-489 (1974).
 RN [4] X-RAY CRYSTALLOGRAPHY (1.3 ANGSTROMS).
 RP X-RAY CRYSTALLOGRAPHY (1.3 ANGSTROMS).
 RX MEDLINE=94194522; PubMed=814525;
 RA Housset D., Habersetzer-Rochat C., Astier J.P., Fontecilla-Camps J.C.;
 RT RT "Crystal structure of toxin II from the scorpion Androctonus australis Hector refined at 1.3 Å resolution.";
 RN [5] X-RAY CRYSTALLOGRAPHY (0.96 ANGSTROMS).
 RP X-RAY CRYSTALLOGRAPHY (0.96 ANGSTROMS).
 RA Smith G.D., Blessing R.H., Ballick S.E., Fontecilla-Camps J.C., Hauptman H.A., Housset D., Langs D.A., Miller R.;
 RT RT "Ab initio structure determination and refinement of a scorpion protein toxin.";
 RT RL
 CC -1- FUNCTION: BINDS TO SODIUM CHANNELS AND INHIBITS THE INACTIVATION OF THE ACTIVATED CHANNELS, THEREBY BLOCKING NEURONAL TRANSMISSION.
 CC -1- SUBCELLULAR LOCATION: Secretd.
 CC -1- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
 CC -1- ALPHA-TOXIN SUBFAMILY.
 CC ---
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).
 CC ---
 DR EMBL: M7704; AAA29949; 1; -.
 DR PIR: A01744; NTSP2A.
 DR PIR: D31444; D34444.
 DR PDB: 1P9X; 26-JAN-95.
 DR PDB: 1AH0; 15-OCT-97.
 DR InterPro: IPR03614; Knot1.
 DR InterPro: IPR02061; Scorpion_toxin.
 DR InterPro: IPR001219; Toxin.
 DR Pfam: PF00537; toxin_3; 1.
 DR PRINTS; PR00384; TOXIN.
 DR SMART: SM00505; Knot1.
 DR Prodom; PD000908; Scorpion_toxin; 1.
 DR SMRT: SM00505; Knot1.
 KW Neurotoxin; Amidation; Sodium channel inhibitor; Signal; 3D-structure.
 FT SIGNAL 1 19
 FT CHAIN 1 19
 FT DISULFID 20 83
 FT DISULFID 31 82
 FT DISULFID 35 82
 FT DISULFID 41 65
 FT DISULFID 45 67
 FT MOD_RES 83 83
 SQ SEQUENCE 85 AA; 95.8 MW; 1740CC6B98363768 CRC64;

Query Match 23.5%; Score 79.5; DB 1; Length 85;
 Best Local Similarity 24.2%; Pred. No. 0.019;
 Matches 16; Conservative 16; Mismatches 27; Indels 7; Gaps 3;
 Qy 1 MKSIAITIVLYAFCILEDGTEAGEFCPFWAGK --CHRHCKSFRRGFCFR -GTEFR 55
 Db 6 MISLAELFVTGVE - SVKDGYLVDDVNTYFGCRNAYCNEFCTLKGESEGYCOWASPYGN 63
 Qy 56 TCYCYR 61
 Db 64 ACYCYR 69

CC --!- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- TISSUE SPECIFICITY: EXPRESSED IN FLOWERS AND IN YOUNG FRUITS.
 CC -|- DEVELOPMENTAL STAGE: ACCUMULATE DURING RIENING.
 CC -|- SIMILARITY: BELONGS TO THE GAMMA-PUROPHONIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; X95730; CAA65046.1; -.
 DR HSSP; P20230; IGPT.
 DR MendeI; 29862; Capn;1533;29862.
 DR InterPro; IPR002118; Gamma-thionin.
 DR InterPro; IPR003614; Knot1.
 DR Pfam; PF003004; Gamma-thionin; 1.
 DR ProDom; PD002594; Gamma-thionin; 1.
 DR SMART; SM00505; Knot1; 1.
 DR PROSITE; PS00940; GAMMA-THIONIN; 1.
 KW Defensin; Fungicide; Signal.
 FT SIGNAL 1 27 BY SIMILARITY.
 CHAIN 28 74 DEFENSIN J1-2.
 FT DISULFID 30 74 BY SIMILARITY.
 FT DISULFID 41 61 BY SIMILARITY.
 FT DISULFID 47 68 BY SIMILARITY.
 FT DISULFID 51 70 BY SIMILARITY.
 SQ SEQUENCE 74 AA; 8249 MW; D92D8F06F39E1552 CRC64;

Query Match 21.1%; Score 71.5%; DB 1; Length 84;
 Best Local Similarity 27.3%; Pred. No. 0.15%;
 Matches 21; Conservative 7; Mismatches 26; Indels 23; Gaps 3;

DR DEFC_AEDAE STANDARD; PRT; 99 AA.
 AC P81603; Q9Y0F0; Q9Y0F0;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Defensin C precursor.
 OS Aedes aegypti (Yellowfever mosquito),
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
 OC Aedes aegypti;
 RN [1] NCBI_TAXID=1159;
 RP SEQUENCE FROM N.A., DEVELOPMENTAL STAGE, INDUCTION, AND SUBCELLULAR LOCATION.
 RC STRAIN=LIVERPOOL; TISSUE=Fat body;
 RX MEDLINE=99124369; PMID=9921719;
 RA Lowenberger C.A.; Smart C.T.; Bulet P., Ferdig M.T., Severson D.W., Hoffmann J.A., Chris-tensen B.M.;
 RT "Insect immunity: molecular cloning, expression, and characterization of cDNAs and genomic DNA encoding three isoforms of insect defensins in Aedes aegypti"; Insect Mol. Biol. 8:107-118(1999).
 RN [2] RP SEQUENCE OF 60-99, AND INDUCTION.
 RC STRAIN=LIVERPOOL;
 RX MEDLINE=95160030; PubMed=7633471;
 RA Lowenberger C., Bulet P., Charlet M., Hetru C., Hodgeman B., Hoffmann J.A.;
 RA Chris-tensen B.M.;
 RT "Insect immunity: isolation of three novel inducible antibacterial defensins from the vector mosquito Aedes aegypti.";
 RL Insect Biochem. Mol. Biol. 25:867-873 (1995).
 RN [3]

CC --!- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -|- SUBCELLULAR LOCATION: EXPRESSED IN FLOWERS AND IN YOUNG FRUITS.
 CC -|- DEVELOPMENTAL STAGE: ACCUMULATE DURING RIENING.
 CC -|- SIMILARITY: BELONGS TO THE GAMMA-PUROPHONIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; D17670; BAA04552.1; -.
 DR HSSP; P10891; IICA.
 DR InterPro; IPR001542; Arthro_defensin.
 DR InterPro; IPR003614; Knot1.
 DR Pfam; PF01077; Arthro_defensin; 1.
 DR PR00271; DEFENSIN.
 DR SMART; SM00505; Knot1; 1.
 DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
 KW Insect immunity; Antibiotic; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT PROPEP 20 41
 FT CHAIN 42 84 TENBCIN 1.
 FT DISULFID 44 75
 FT DISULFID 61 81
 FT DISULFID 65 83
 SQ SEQUENCE 84 AA; 9176 MW; 0367C5070468BE60 CRC64;

Query Match 21.1%; Score 71.5%; DB 1; Length 84;
 Best Local Similarity 27.3%; Pred. No. 0.15%;
 Matches 21; Conservative 7; Mismatches 26; Indels 23; Gaps 3;

QY 4 IAIIFTIVLAFCILEDGIDGI-EVAG------FGCPNAGKCHRICKSI 42
 DEFC_AEDAE STANDARD; PRT; 99 AA.
 Db 9 VACFFIQLIAAFPLIEAAATAEELEQGBHTRVKRTCDILSYBAKGVNLDAACAAICLFR 68

QY 43 RRGGEFCRGTERTTCYC 59
 Db 69 GRSGGYCNG--KRVCYC 83

RESULT 14
 DEFC_AEDAE
 ID DEFC_AEDAE STANDARD; PRT; 99 AA.
 AC P81603; Q9Y0F0; Q9Y0F0;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Defensin C precursor.
 OS Aedes aegypti (Yellowfever mosquito),
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
 OC Aedes aegypti;
 RN [1] NCBI_TAXID=1159;
 RP SEQUENCE FROM N.A., DEVELOPMENTAL STAGE, INDUCTION, AND SUBCELLULAR LOCATION.
 RC STRAIN=LIVERPOOL; TISSUE=Fat body;
 RX MEDLINE=99124369; PMID=9921719;
 RA Lowenberger C.A.; Smart C.T.; Bulet P., Ferdig M.T., Severson D.W., Hoffmann J.A., Chris-tensen B.M.;
 RT "Insect immunity: molecular cloning, expression, and characterization of cDNAs and genomic DNA encoding three isoforms of insect defensins in Aedes aegypti"; Insect Mol. Biol. 8:107-118(1999).
 RN [2] RP SEQUENCE OF 60-99, AND INDUCTION.
 RC STRAIN=LIVERPOOL;
 RX MEDLINE=95160030; PubMed=7633471;
 RA Lowenberger C., Bulet P., Charlet M., Hetru C., Hodgeman B., Hoffmann J.A.;
 RA Chris-tensen B.M.;
 RT "Insect immunity: isolation of three novel inducible antibacterial defensins from the vector mosquito Aedes aegypti.";
 RL Insect Biochem. Mol. Biol. 25:867-873 (1995).
 RN [3]

CC --!- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -|- SUBCELLULAR LOCATION: EXPRESSED IN FLOWERS AND IN YOUNG FRUITS.
 CC -|- DEVELOPMENTAL STAGE: ACCUMULATE DURING RIENING.
 CC -|- SIMILARITY: BELONGS TO THE GAMMA-PUROPHONIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; X95730; CAA65046.1; -.
 DR HSSP; P20230; IGPT.
 DR MendeI; 29862; Capn;1533;29862.
 DR InterPro; IPR002118; Gamma-thionin.
 DR InterPro; IPR003614; Knot1.
 DR Pfam; PF003004; Gamma-thionin; 1.
 DR ProDom; PD002594; Gamma-thionin; 1.
 DR SMART; SM00505; Knot1; 1.
 DR PROSITE; PS00940; GAMMA-THIONIN; 1.
 KW Defensin; Fungicide; Signal.
 FT SIGNAL 1 27 BY SIMILARITY.
 CHAIN 28 74 DEFENSIN J1-2.
 FT DISULFID 30 74 BY SIMILARITY.
 FT DISULFID 41 61 BY SIMILARITY.
 FT DISULFID 47 68 BY SIMILARITY.
 FT DISULFID 51 70 BY SIMILARITY.
 SQ SEQUENCE 74 AA; 8249 MW; D92D8F06F39E1552 CRC64;

Query Match 21.1%; Score 71.5%; DB 1; Length 74;
 Best Local Similarity 30.0%; Pred. No. 0.13%;
 Matches 21; Conservative 8; Mismatches 28; Indels 13; Gaps 3;

QY 2 KSIAIIFTIVLAFCILEDGIDGI-EVAG------FNAGKCHRICKSIRRGGFCRG 51
 Db 6 KVATIIFTLMMM-LVFTAGMVAEARTCESOSHFRKGLCFKNSNGSVCHTEFGNGHCRG 63

QY 52 TFRRTTCYCVR 61
 Db 64 -FRRRCFCTR 72

RESULT 13
 DEFI_TENMO STANDARD; PRT; 84 AA.
 ID DEFI_TENMO STANDARD; PRT; 84 AA.
 AC Q27023; 35, Created
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE Tenecin 1 precursor.
 OS Tenecin 1 precursor (Yellow mealworm).
 OC Tenebrio molitor (Yellow mealworm).
 OC Pterygota; Metzoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Cucujiformia; Tenebrionidae; Tenebrio.
 RN [1] SEQUENCE FROM N.A., SEQUENCE OF 49-59, AND DISULFIDE BONDS.
 RP Purification and molecular cloning of cDNA for an inducible protein from larva of the coleopteran, *Tenebrio molitor*.
 RP Purification and molecular cloning of cDNA for an inducible protein from larva of the coleopteran, *Tenebrio molitor*.
 RL J. Biochem. 116:53-58 (1994).
 CC -|- FUNCTION: BACTERICIDAL PROTEIN PRODUCED IN RESPONSE TO INJURY . IT IS CYTOTOXIC TO GRAM-POSITIVE BACTERIA.
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.

SEQUENCE OF 60-99.
STRAIN=REFM; PubMed=1568225;
MEDLINE=96047965;

RC Chalk R, Albuquerque C.M., Ham P.J., Townson H.;
Full sequence and characterization of two insect defensins: immune peptides from the mosquito Aedes aegypti.;
Proc. R. Soc. Lond. B, Biol. Sci. 261:217-221(1995).

RX -1- FUNCTION: ANTBACTERIAL PEPTIDE MOSTLY AGAINST GRAM-POSITIVE BACTERIA.

RT -1- SUBCELLULAR LOCATION: SECRETED IN THE HEMOLYMPH. EXPRESSED 30 MINUTES AFTER INFECTION AND REMAINED PRESENT THROUGH TO 21 DAYS. EXPRESSED IN WHITE OR CALLOUS PUPAE DURING METAMORPHOSIS, BUT NO EXPRESSION WAS SEEN IN LARVAE.

RL -1- INDUCTION: By bacterial infection.

CC -1- MORPHISM: THERE ARE TWO DEFENSIN C ISOFORMS, C1 (SHOWN HERE) AND C2.

CC -1- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF156092; AAD40116.2;
DR EMBL; AF156093; AAD40117.2; .
DR HSSP; P10891; IICCA.
DR InterPro; IPRO01542; Arthro_defensin.
DR InterPro; IPRO03614; Knot1.
DR Pfam; PF002594; Gamma-thionin; 1.
DR PRINTS; PRO00271; DEFENSIN.
. DR SMART; SM00505; Knot1; 1.
DR PROSITE; PS00940; GAMMA-THONIN; 1.
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS_1.
KW Insect immunity; Antibiotic; Defensin; Signal; Polymorphism.
FT SIGNAL 1 23 POTENTIAL.
FT PROTEP 24 59 DEFENSIN C.
FT CHAIN 60 99 BY SIMILARITY.
FT DISULFID 62 89 BY SIMILARITY.
FT DISULFID 75 95 BY SIMILARITY.
FT DISULFID 79 97 BY SIMILARITY.
FT VARIANT 26 26 G -> E (IN ISOFORM C2).
FT VARIANT 34 34 P -> S (IN ISOFORM C2).
FT CONFLICT 83 83 R -> G (IN REF. 2 AND 3).
FT CONFLICT 91 91 A -> S (IN REF. 2 AND 3).
SEQUENCE 99 AA; 10696 MW;

[1] SEQUENCE FROM N.A.

RN RP TISSUE=Flower;
RC Urdangarin M.C., de la Canal L.;
RA "A defensin gene expressed in sunflower inflorescence.";
RT "A defensin gene expressed in sunflower inflorescence.";
RL Plant Physiol. Biochem. 38:253-258(2000).
CC -1- FUNCTION: MAY PLAY A PROTECTIVE ROLE IN FLOWERS BY PROTECTING THE REPRODUCTIVE ORGANS FROM POTENTIAL PATHOGEN ATTACK.

CC -1- SUBCELLULAR LOCATION: CELL WALL OR VACUOLE (POTENTIAL).

CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN FLOWERS AND TO A LESSER EXTENT IN LEAVES. LOWER LEVELS IN HYPOCOTYLUS. NO EXPRESSION IN ROOTS AND COOTYLEDONS.

CC -1- DEVELOPMENTAL STAGE: EXPRESSED PROGRESSIVELY DURING FLOWER DEVELOPMENTAL STAGE, REACHING THE HIGHEST LEVEL IN THE MATURE FERTILIZED FLOWER STAGE.

CC -1- SIMILARITY: BELONGS TO THE GAMMA-PYROTHONIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF178634; AAF72042.1; .
DR EMBL; AF141131; AAF6591.1; .
DR InterPro; IPRO02118; Gamma-thionin.
DR InterPro; IPRO03614; Knot1.
DR Pfam; PF00304; Gamma-thionin; 1.
DR PRINTS; PRO00288; PYROTHONIN.
DR ProDom; PD002594; Gamma-thionin; 1.
DR SMART; SM00505; Knot1; 1.
DR PROSITE; PS00940; GAMMA-THONIN; 1.
KW Thionin; Plant toxin; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 78 FLOWER SPECIFIC GAMMA-THONIN.
FT DISULFID 34 78 BY SIMILARITY.
FT DISULFID 45 65 BY SIMILARITY.
FT DISULFID 51 72 BY SIMILARITY.
FT DISULFID 55 74 BY SIMILARITY.
SQ SEQUENCE 78 AA; 8627 MW; IAD6A98582149A2D CRC64;

Query Match Best Local Similarity 36.0%; Score 70.5%; DB 1; Length 78; Matches 27; Conservative 6; Mismatches 25; Indels 17; Gaps 5;

Qy 1 MKSIAIF--IVLVAFCL--EDG--IVEAGF-----GCPNAGKCHRKSKIRR 44
Db 1 MKSSMMRFAALLVVMLLANEMGGELVARTCESQSHFKGTCLSDFNCVNHCSRF 60

Query Match Best Local Similarity 41.7%; Score 71; DB 1; Length 99; Matches 15; Conservative 5; Mismatches 12; Indels 4; Gaps 2;

Qy 24 AGFGCPFNAGKCHRHKCSIRRGGFCRGTFRTTCVC 59
Db 66 SGFGYGDSAA-CAAHCIAFRNRGGYCNAA-KKVVCV 97

RESULT 15
THGF_HELAN STANDARD PRT; 78 AA.

Best Local Similarity 41.7%; Pred. No. 0.19; Score 71; DB 1; Length 99; Matches 15; Conservative 5; Mismatches 12; Indels 4; Gaps 2;

Qy 24 AGFGCPFNAGKCHRHKCSIRRGGFCRGTFRTTCVC 59
Db 66 SGFGYGDSAA-CAAHCIAFRNRGGYCNAA-KKVVCV 97

Search completed: September 17, 2002, 15:45:49
Job time: 125 sec

Result No.	Score	Query	Match	Length	DB	ID	Description
1	177	52.2	38	2	JN0613		defensin 4 K - scor
2	133.5	39.4	38	2	S27242		defensin - blue da
3	128	37.8	38	2	S74088		defensin - Mediter
4	80.5	23.7	43	2	B41711		defensin B - beetl
5	79.5	23.5	85	1	NTSR2A		neurotoxin II prec
6	74.5	22.0	43	2	C41711		defensin C - beetl
7	72.5	21.4	43	2	JC2554		holotrichin 1 - Hol
8	71.5	21.1	84	2	JX0332		teneclin 1 precursor
9	69	20.5	65	1	NTSRB		neurotoxin XI - sc
10	69	20.4	75	2	S11156		pSAS10 protein - c
11	68.5	20.2	72	2	T06599		disease resistance
12	68	20.1	94	2	A31792		sapecin precursor
13	67.5	19.9	79	2	T06381		protease inhibitor
14	66.5	19.6	64	1	NTSRB		neurotoxin III - s
15	66.5	19.6	86	1	JN0671		Na+-channel blocki
16	66	19.5	94	2	S12558		defensin A precurs
17	65.5	19.3	64	1	NTSR5L		neurotoxin V - Egy
18	65.5	19.3	64	1	NTSR5M		neurotoxin V - sco
19	64.5	19.0	798	2	A40526		integrin beta-7 ch
20	63.5	18.7	77	2	B8443		protease inhibitor
21	63.5	18.7	80	2	T10823		antifungal protein
22	63.5	18.7	768	2	B41029		integrin beta-8 ch
23	63.5	18.7	769	2	A41029		integrin beta-8 ch
24	63	18.6	90	2	B84867		probable trypsin i
25	62	18.3	990	2	151618		nucleolar phosphop
26	61.5	18.1	77	2	S30578		protease inhibitor
27	61.5	18.1	80	2	T10176		antifungal protein
28	61.5	18.1	154	2	T17816		hypothetical prote
29	61	18.0	40	2	JU0225		sapecin C - flesh

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:42:49 ; Search time 18:35 Seconds
(without alignments)
319.425 Million cell updates/sec

Title: US-09-829-481-4
Perfect score: 339
Sequence: 1 MKSIAIIIFIVLVAFCLLEDG.....IRRGFCRGTFRTTCVYR 61

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:
1: Pirl;*
2: Pir2;*
3: Pir3;*
4: Pirl;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

RESULT	Score	Query	Match	Length	DB	ID	Description
1	JN0613	defensin 4K - scorpion (Leiurus quinquestriatus)					N: Alternative names: antibacterial 4K peptide
							C: Species: Leiurus quinquestriatus Hebraeus
							C: Accession: JN0613
							R: Cocianich, S.; Goffon, M.; Bontems, F.; Bouet, P.; Menez, A.; Hoffmann Biophys. Res. Commun. 194, 17-22, 1993
							A: Title: Purification and characterization of a scorpion defensin, a 4kDa antibacteri
							A: Reference number: JN0613; MUID: 93326112
							A: Accession: JN0613
							A: Molecule type: protein
							A: Residues: 1-38 <COC>
							A: Note: this protein is similar to scorpion toxins and insect defensins
							Query Match 52.2%; Score 177; DB 2; Length 38;
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Matches 27; Conservative
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61
							Best Local Similarity 73.0%; Pred. No. 5.4e-13; 3; Mismatches 7; Indels 0; Gaps 0;
							Db 1 GFGCPILNQAGCHRCSIRRGGYCAGFFKQNCCTCR 37
							Query Match 25 GFGCPIFNAGKCHRCKSIRRGGFCRTTCVYR 61

S74088 Defensin - Mediterranean mussel (Mediterranean mussel)
 C;Species: *Mitius galloprovincialis* (Mediterranean mussel)
 C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 04-Feb-2000
 C;Accession: S74088
 R;Hubert, F.; Noel, T.; Roch, P.
 Bur. J. Biochem. 240, 302-306, 1996
 A;Title: A member of the arthropod defensin family from edible Mediterranean mussels (Mytilus galloprovincialis) from the scorpion Androctonus australis hector
 A;Reference number: S74088; MUID:97025339
 A;Molecule type: protein
 A;Residues: 1-38 <BUP>
 A;Experimental source: hemolymph
 C;Keywords: antibiotic
 F;4-25,10-33,14-35,21-38/Disulfide bonds: #status predicted

Query Match	37.8%;	Score 128;	DB 2;	Length 38;
Best Local Similarity	63.3%;	Pred. No. 1.5e-07;		
Matches 24;	Conservative	2;	Gaps 10;	Indels 2;
Oy	25 GFGCPNAGGCHRCKSTRR-GGFORGTPTTCYCYR 61			
Db	1 GFGCP-NNYQCHRHKCKSTPGRCGGYCGGXHFLRCTCYR 37			

RESULT 4

B41711	defensin B - beetle (Zophobas atratus)
C;Species: Zophobas atratus	
C;Accession: B41711	
Oy	24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 03-May-1996
Db	

RESULT 5

B41711	defensin C - beetle (Zophobas atratus)
C;Species: Zophobas atratus	
C;Accession: B41711	
Oy	24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 03-May-1996
Db	

RESULT 5

N7NSR2A	neurotoxin II precursor [validated] - Sahara scorpion
N;Alternate names: mammalian neurotoxin (Clone pCD402)	
C;Species: Androctonus australis (Sahara scorpion)	
C;Date: 24-Apr-1984 #sequence_revision 12-Apr-1996 #text_change 15-Sep-2000	
C;Accession: D34444; A01744	
R;Boujat, P.E.; Rochat, H.; Smith, L.A.	
J. Biol. Chem. 264, 19259-19265, 1989	
A;Title: Precursors of Androctonus australis scorpion neurotoxins. Structures of precursors	
A;Reference number: A34444; MUID:90037062	
A;Molecule type: mRNA	
A;Residues: 1-85 <BUP>	
A;Cross-references: GB:M27704; NID:9161144; PID:AAA29949.1; GB:J05102	
A;Experimental source: clone pCD402	
R;Rochat, H.; Rochat, C.; Sampieri, F.; Miranda, F.; Lissitzky, S.	
Bur. J. Biochem. 28, 381-388, 1972	
A;Title: The amino-acid sequence of neurotoxin II of Androctonus australis Hector.	
A;Reference number: A91198; MUID:73025153	
A;Accession: A01744	

A;Molecule type: protein
 A;Residues: 20-83 <ROC>
 R;Kopeyan, C.; Martinez, G.; Lissitzky, S.; Miranda, F.; Rochat, H.
 Eur. J. Biochem. 47, 483-489, 1974
 A;Title: Disulfide bonds of toxin II of the scorpion Androctonus australis hector
 A;Reference number: A91225; MUID:75057756
 A;Contents: annotation; disulfide bonds
 R;Houset, D.; Habersetzer-Rochat, C.; Astier, J.P.; Fontecilla-Camps, J.C.
 J. Mol. Biol. 238, 88-103, 1994
 A;Title: Crystal structure of toxin II from the scorpion Androctonus australis hector
 A;Reference number: A58108; MUID:94194522
 A;Contents: annotation; X-ray crystallography, 1.3 angstroms, residues 20-83
 R;Fontecilla-Camps, J.C.; Houset, D.
 Submitted to the Brookhaven Protein Data Bank, September 1994
 A;Reference number: A67142; PDB:1LPTX
 A;Contents: annotation; X-ray crystallography, 1.3 angstroms, residues 20-83
 C;Superfamily: scorpion neurotoxin
 C;Keywords: amidated carboxyl end; neurotoxin; venom
 F;1-19/Domain: signal sequence #status Predicted <SIG>
 F;20-83/Product: neurotoxin II #status experimental <MAT>
 F;31-82,35-55,41-65,45-67/Disulfide bonds: #status experimental
 F;83/Nodified site: amidated carboxyl end (His) (amide in mature form from following
 Query Match
 23.5%; | Score 79.5; | DB 1; | Length 85; || Best Local Similarity | 24.2%; | Pred. No. 0.062; | | |
Matches 16;	Conservative	16;	Gaps 16;	Indels 7;
Oy	1 MKSIAITIFIVLVAFTILEDGIVEAGFGCPFNAGK --CHRHCKSIRRGGFCR -GTFR 55			
Db	6 MISLAUFLFVGE-- -SVKGIVDDVNCTYFCGRNAYCNEECTKLKGESGYCQWASPYGN 63			
Query Match	23.5%;	Score 79.5;	DB 1;	Length 85;
Best Local Similarity	24.2%;	Pred. No. 0.062;		
Matches 16;	Conservative	16;	Gaps 16;	Indels 7;
Oy	56 TCVCYR 61			
Db	64 ACYCYK 69			
RESULT 6				
C41711	defensin C - beetle (Zophobas atratus)			
C;Species: Zophobas atratus				
C;Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 03-May-1996				
C;Accession: C41711				
R;Buler, P.; Cocianich, S.; Dimarco, J.L.; Lambert, J.; Reichhart, J.M.; Hoffmann, D.; J. Biol. Chem. 266, 24520-24525, 1991				
A;Title: Insect immunity. Isolation from a coleopteran insect of a novel inducible anti-pathogen defense protein.				
A;Reference number: A41711; MUID:92105112				
A;Accession: B41711				
A;Status: preliminary				
A;Molecule type: protein				
A;Residues: 1-43 <BU1>				
C;Superfamily: sapecin precursor				
Query Match	23.7%;	Score 80.5;	DB 2;	Length 43;
Best Local Similarity	39.3%;	Pred. No. 0.03;		
Matches 15;	Conservative	6;	Gaps 12;	Indels 5;
Oy	25 GF---GCPFNAGKCHRCKSTRRGGFCRGTFR 59			
Db	7 GFEIAGTKLNAAACGAHCLALGRRGGYCNS -KSVCVC 42			
RESULT 5				
N7NSR2A	neurotoxin II precursor [validated] - Sahara scorpion			
N;Alternate names: mammalian neurotoxin (Clone pCD402)				
C;Species: Androctonus australis (Sahara scorpion)				
C;Date: 24-Apr-1984 #sequence_revision 12-Apr-1996 #text_change 15-Sep-2000				
C;Accession: D34444; A01744				
R;Boujat, P.E.; Rochat, H.; Smith, L.A.				
J. Biol. Chem. 264, 19259-19265, 1989				
A;Title: Precursors of Androctonus australis scorpion neurotoxins. Structures of precursors				
A;Reference number: A34444; MUID:90037062				
A;Molecule type: mRNA				
A;Residues: 1-85 <BO1>				
A;Cross-references: GB:M27704; NID:9161144; PID:AAA29949.1; GB:J05102				
A;Experimental source: clone pCD402				
R;Rochat, H.; Rochat, C.; Sampieri, F.; Miranda, F.; Lissitzky, S.				
Bur. J. Biochem. 28, 381-388, 1972				
A;Title: The amino-acid sequence of neurotoxin II of Androctonus australis Hector.				
A;Reference number: A91198; MUID:73025153				
A;Accession: A01744				
RESULT 7				
JC2554	holotrichin 1 - Holotrichia diomphalia			
C;Species: Holotrichia diomphalia				
C;Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 17-Mar-1999				
C;Accession: JC2554				
R;Lee, S.Y.; Moon, H.-J.; Kawabata, S.; Kurata, S.; Natori, S.; Lee, B.L.				
Biol. Pharm. Bull. 18, 457-459, 1995				
A;Title: A sapecin homologue of Holotrichia diomphalia: Purification, sequencing and				
A;Reference number: JC2554; MUID:96031330				

A;Accession: JC2554
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-43 <LEE>
C;Superfamily: sapecin precursor

Query Match 21.4%; Score 72.5; DB 2; Length 43;
Best Local Similarity 39.4%; Pred. No. 0.23; Indels 3; Gaps 1;
Matches 13; Conservative 5; Mismatches 12; Indels 3; Gaps 1;

Qy 27 GCPFNAGKCHRICKSIRRGGFCRGTFRTTCVC 59
Db 12 GIALNDSCAAHCLAMRKGGSK--QGVCVC 41

RESULT 8

tenecin 1 precursor - yellow mealworm
C;Accession: JX0332
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 20-Jun-2000
R;Moon, H.J.; Lee, S.Y.; Kurata, S.; Natori, S.; Lee, B.L.
J. Biochem. 116: 53-58, 1994
A;Title: Purification and molecular cloning of cDNA for an inducible antibacterial protein
A;Reference number: JX0332; MUID:95096025
A;Molecule type: mRNA
A;Residues: 1-84 <HOO>
A;Cross-references: GB:D17670; NID:q1235940; PIDN:BA004552.1; PID:q1235941
A;Experimental source: larva
C;Comment: Tenecin 1 is an antibacterial protein induced in the hemolymph of larvae of t
C;Superfamily: sapecin precursor
F;1-11/Domain: signal sequence #status predicted <SIG>
F;42-84/Product: tenecin 1 #status predicted <MAT>
F;44-75,61-81,65-83/Disulfide bonds: #status predicted

Query Match 21.1%; Score 71.5; DB 2; Length 84;
Best Local Similarity 27.3%; Pred. No. 0.48; Indels 26; Gaps 3;
Matches 21; Conservative 7; Mismatches 26; Indels 23; Gaps 3;

Qy 4 IAIPIIVLVAFCFILEDGLI----VEAG-----FGCPFNAGKCHRICKS 42
Db 9 VACFTIQLAAFPFLPEATAEITQGEHIRVKRVTCDILSVEAKVKAINDAAAHCLFR 68

Qy 43 RRRGGFCRGTFRTTCVC 59
Db 69 GRSGGYCNG -KRVCVC 83

RESULT 9

NTSREB neurotoxin XI - scorpion (Buthus occitanus)
C;Species: Buthus occitanus tunetanus
C;Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 23-Aug-1996
C;Accession: A01746
R;Martin, M.F.; Rochat, H.
Toxicon 22, 279-291, 1984
A;Title: Purification of thirteen toxins active on mice from the venom of the North African
A;Reference number: A94316; MUID:84224414
A;Accession: A01746
A;Molecule type: protein
A;Residues: 1-65 <MAR>
C;Keywords: scorpion neurotoxin
C;Keywords: neurotoxin; venom
F;12-63,16-36,22-46,26-48/Disulfide bonds: #status predicted

Query Match 20.5%; Score 69.5; DB 1; Length 65;
Best Local Similarity 26.0%; Pred. No. 0.66; Mismatches 10; Indels 5; Gaps 2;

Qy 17 LEDGIVIEAGFGCPFNAGK - -CHRICKSIRRGGFCR - -GTFRTTCVCYR 61
Db 1 LKDGYIVDDRNCTYFCGTNAYCNFECVKLRGESYCYQWGRYGNACWCYR 50

RESULT 10

S11156 protein - cowpea
PSAS10 protein - cowpea
C;Species: Vigna unguiculata (cowpea)
C;Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jan-2000
B;Ishibashi, N.; Yamada, D.; Minamikawa, T.
Plant Mol. Biol. 15, 59-64, 1990
A;Title: Stored mRNA in cotyledons of Vigna unguiculata seeds: nucleotide sequence of
A;Reference number: S11156; MUID:91355865
A;Accession: S11156
A;Molecule type: mRNA
A;Status: preliminary
A;Cross-references: EMBL:X16877; NID:q22075; PID:q22076
C;Superfamily: gamma-thionin

Query Match 20.4%; Score 69; DB 2; Length 75;
Best Local Similarity 33.3%; Pred. No. 0.84; Mismatches 6; Indels 14; Gaps 4;

Qy 2 KSIAT-ILFIVLVAFCILE-----DGIVAEAGFGCPFNAGKCHRICKSIRRGGFC 49
Db 4 KTAGICFLFLVLFVQAEEVVVQSEARTCENLVDTYRGPFCFTGSCUDHCKRNKEHLLSGRC 63

Qy 50 RGTFRPTTCVCYR 61
Db 64 RDDVR - CWCTR 73

RESULT 11

T06599 disease resistance response protein 230 - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jun-2000
C;Accession: T06599
R;Chiang, C.C.; Hadwiger, L.A.
Mol. Plant Microbe Interact. 4, 324-331, 1991
A;Title: The Fusarium solani-induced expression of a pea gene family encoding high cy
A;Reference number: 215787; MUID:92190628
A;Accession: T06599
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-72 <CHT>
A;Cross-references: EMBL:L01578; PIDN:AAA79117.1; PID:q169074
A;Experimental source: cv. Alaska
C;Genetics:
A;Gene: DRR230-a
C;Superfamily: gamma-thionin

Query Match 20.2%; Score 68.5; DB 2; Length 72;
Best Local Similarity 28.4%; Pred. No. 0.92; Mismatches 11; Indels 13; Gaps 3;

Qy 2 KSTATIIFVLVAFCILEDGVAGFGCPFNAGK-----CHRICKSIRRGG---GFCCR 50
Db 4 KSIACLSFLLVDFVAOEIVVSANTENLAGSYKGVCFGGDHRCT - QEGANISGRCR 61

Qy 51 GTFRTTC 57
Db 62 DDFRWC 68

RESULT 12

A31792 sapecin precursor - flesh fly (Sarcophaga peregrina)

N:Alternate names: antibacterial protein
 C:Species: Sarcophaga pereirae
 C:Accession: A31792; A31791
 R:Matsuura, K.; Natori, S.
 J. Biol. Chem. 263, 17117-17121, 1988

A:Title: Molecular cloning of cDNA for sapecin and unique expression of the sapecin gene
 A:Reference number: A31792; MUID:89034216
 A:Accession: A31792

A: Molecule type: mRNA
 A: Residues: 1-94 <MAT>
 A: Cross-references: GB-J04053; NID:9161266; PIDN:AAA29984.1; PID:g161267

R:Matsuura, K.; Natori, S.
 J. Biol. Chem. 263, 17112-17116, 1988
 A:Title: purification of three antibacterial proteins from the culture medium of NIH-Sap
 A:Reference number: A92687; MUID:89034215
 A:Accession: A31791

A: Molecule type: protein
 A: Residues: 55-94 <MA2>
 C:Superfamily: sapecin precursor

Query Match 20.1% Score 68; DB 2; Length 94;
 Best Local Similarity 39.4%; Pred. No. 1..3.;
 Matches 13; Conservative 2; Mismatches 16; Indels 2; Gaps 1;

QY 27 GCPFNAGKCHRHCKSIRRGGFCRGTFRITCVC 59
 DB 62 GTINGHSACAAHCLLRLGNRGYCNG -KAVCVC 92

RESULT 13
 protease inhibitor - soybean
 C:Species: Glycine max (soybean)

C:Accession: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jun-2000
 R:Accession: T06381
 R:Accession: T06381

R:Description: Isolation and expression of a drought-induced protease inhibitor from soy
 A:Reference number: 215640
 A:Accession: T06381

A:Status: preliminary; translated from GB/EMBL/DBDJ

A: Molecule type: mRNA

A:Cross-references: EMBL:U12150; NID:9533691; PIDN:ACC97524.1; PID:9533692

A:Experimental source: strain Essex; leaf
 C:Superfamily: gamma-thionin

Query Match 26.1% Score 67.5; DB 2; Length 79;
 Best Local Similarity 26.1%; Pred. No. 1..3.;
 Matches 18; Conservative 13; Mismatches 27; Indels 11; Gaps 2;

QY 1 MKSTIALIFTVLAFILEDGVIEAGFGCPFNAGK -----CHRICKSIRRGGFCR 50
 DB 8 VSTICVLLVLVATEMMGPTMVAEARTCESQSHEREKGPCLSDTNCGSVCRTERFTGSHCR 67

QY 51 GTFRTTCVC 59
 DB 68 G-FRRRCFC 75

RESULT 14
 NTSR3B
 neurotoxin III - scorpion (Butthus occitanus)
 C:Species: Butthus occitanus tunetanus
 C:Accession: A01745
 R:Vargas, O.; Gregoire, J.; Martin, M.F.; Bechis, G.; Rochat, H.
 Toxicon 20, 79, 1982
 A:Title: Neurotoxins from the venoms of two scorpions: Butthus occitanus tunetanus and Bu

N:Alternate names: antibacterial protein
 C:Species: Sarcophaga pereirae
 C:Accession: A01745
 A: Molecule type: protein
 A: Residues: 1-64 <VAR>
 C:Superfamily: scorpion neurotoxin
 C:Keywords: blocked carboxyl end; neurotoxin; venom
 F:12-63 16-36 22-46 26-48/Dsulfide bonds; #status predicted
 F:64/Modified site: blocked carboxyl end (Asn) (probably amidated) #status experiment

Query Match 19.6% Score 66.5; DB 1; Length 64;
 Best Local Similarity 22.0%; Pred. No. 1..4;
 Matches 11; Conservative 12; Mismatches 22; Indels 5; Gaps 2;
 QY 17 LEDGIVEAGFGCPFNAGK -----CHRICKSIRRGGFCR -GTFRTTCVCYR 61
 DB 1 VKBDGYIVDDRNCITYFCGFRNAYCNEECTKLKGESGYCOWASPGNACYC 50

RESULT 15
 JN0671
 Na+-channel-blocking toxin (clone cnq1IV) precursor - scorpion (Centruroides noxioides)
 C:Species: Centruroides noxioides

C:Accession: JN0671
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 16-Jul-1999

R:Baccerini, B.; Vazquez, A.; Garcia, C.; Corona, M.; Bolivar, F.; Possani, L.D.
 Gene 128, 165-171, 1993
 A:Title: Cloning and characterization of cDNAs that code for Na+-channel-blocking tox

A:Reference number: JN0669; MUID:93232983
 A:Accession: JN0671
 A: Molecule type: mRNA
 A: Residues: 1-86 <BECA>
 A:Cross-references: GB-L05062; NID:9304570; PIDN:AAA28287.1; PID:9304571

A:Experimental source: venom gland
 C:Superfamily: scorpion neurotoxin
 C:Keywords: toxin
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-86/Product: Na+-channel-blocking toxin (clone cnq1IV) #status predicted <NaT>

Query Match 19.6% Score 66.5; DB 2; Length 86;
 Best Local Similarity 33.3%; Pred. No. 1..7;
 Matches 23; Conservative 7; Mismatches 26; Indels 13; Gaps 4;

QY 1 MKSTAIIT -FIVLVAFILEDGVIEAGFGCPFNAGK -----CHRICKSIRRGG -GFCR 50
 DB 1 MSLLITIACLVLYLIGTVWAKDSYLVDVKGRKRNCTYLGENDYCNRECKMKHRGGSYGYC 60

QY 51 GTFRTTCVC 59
 DB 61 G ---FGCYC 66

Search completed: September 17, 2002, 15:45:00
 Job time: 131 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model
Run on: September 17, 2002, 15:43:24 : Search time 25.63 Seconds
(Without alignments)
411.732 Million cell updates/sec

Title: US-09-829-481-4
Perfect score: 339
Sequence: 1 MKSIAIIIFIVLVAFCTILEDG.....IRRGGFCRGHFRTCVCYR 61

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19_*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_invertebrate:*
- 5: sp_human:*
- 6: sp_invertebrate:*
- 7: sp_mammal:*
- 8: sp_mhc:*
- 9: sp_organelle:*
- 10: sp_plant:*
- 11: sp_rhade:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archeap:*

ALIGNMENTS

Result No.	Score	Query Match	Length	DB ID	Description
1	154.5	45.6	81	09Y0A9	Q9Y0A9 mytilus gal
2	152	44.8	82	5	Q96049 oryctes rhi
3	146.5	43.2	73	5	Q96049 mytilus gal
4	140	41.3	73	5	Q96049 ornithodoros
5	108	31.9	57	5	Q96049 ornithodoros
6	90	26.5	91	5	Q96049 oryctes rhi
7	82.5	24.3	85	5	Q96049 oryctes rhi
8	79.5	23.5	79	5	Q96049 oryctes rhi
9	77	22.7	57	5	Q96049 ornithodoros
10	75.5	22.3	85	5	Q96049 ornithodoros
11	75.5	22.3	85	5	Q96049 ornithodoros
12	71.5	21.1	85	5	Q96049 oryctes rhi
13	70	20.6	102	5	Q96049 ornithodoros
14	70	20.6	103	5	Q96049 ornithodoros
15	69.5	20.5	72	10	Q96049 piusum sativ
16	68.5	20.2	85	5	Q96049 piusum sativ

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Query Length	DB ID	Description
1	154.5	45.6	81	09Y0A9
2	152	44.8	82	5
3	146.5	43.2	73	5
4	140	41.3	73	5
5	108	31.9	57	5
6	90	26.5	91	5
7	82.5	24.3	85	5
8	79.5	23.5	79	5
9	77	22.7	57	5
10	75.5	22.3	85	5
11	75.5	22.3	85	5
12	71.5	21.1	85	5
13	70	20.6	102	5
14	70	20.6	103	5
15	69.5	20.5	72	10
16	68.5	20.2	85	5

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Query Length	DB ID	Description
1	154.5	45.6	81	09Y0A9
2	152	44.8	82	5
3	146.5	43.2	73	5
4	140	41.3	73	5
5	108	31.9	57	5
6	90	26.5	91	5
7	82.5	24.3	85	5
8	79.5	23.5	79	5
9	77	22.7	57	5
10	75.5	22.3	85	5
11	75.5	22.3	85	5
12	71.5	21.1	85	5
13	70	20.6	102	5
14	70	20.6	103	5
15	69.5	20.5	72	10
16	68.5	20.2	85	5

RESULT 1

Q9Y0A9	PRELIMINARY;	PRT;	81 AA.
TD	Q9Y0A9;		
AC	Q9Y0A9;		
DT	01-NOV-1999 (TREMBLrel. 12, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	DEFENSIN MGD-2 PRECURSOR.		
OS	Mytilus galloprovincialis (Mediterranean mussel).		
OC	Bivalvota; Metacoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;		
OC	Mytiloidea; Mytilidae; Mytilus.		
OX	NCBI_TAXID=29158;		
RN	11		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=HEMOCYTE;		
RC	MEDLINE=20033584; PubMed=10564642;		
RA	Milta G., Vandebulcke F., Hubert F., Roch P.;		
RA	"Mussel defensins are synthesized and processed in granulocytes then released into the plasma after bacterial challenge."		
RL	J. Cell Sci. 112:4233-24 (1999).		
CC	- - MISCELLANEOUS: THE PI OF THIS PROTEIN IS 8.92.		
CC	- - SUBCELLULAR LOCATION: SECRETED.		
CC	- - SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.		
CC	- - SIMILARITY: BELOWS TO THE ARTHROPOD DEFENSIN.		
DR	DR AF162338; AAD45118.1;		
DR	InterPro: IPR01542; Arthro_defensin.		
DR	InterPro: IPR005614; Knot1.		
DR	PFAM: PF00197; Arthro_defensin; 1.		
SMART	SM00505; Knot1; 1.		
PROSITE	PS00425; ARTHROPOD_DEFENSINS; UNKNOWN_1.		
DR	Defensin; Antibiotic; Fungicide; Annotation; Signal.		
KW	KW SIGNAL 1 21	POTENTIAL.	
DR	PROPEP 22 60	POTENTIAL.	
FT	CHAIN 61 81	DEFENSIN IN MGDB-2.	
FT	MOD_RES 60 60	AMIDATION (POTENTIAL).	
FT	DISULFID 25 46	BY SIMILARITY.	
FT	DISULFID 31 54	BY SIMILARITY.	
FT	DISULFID 35 56	BY SIMILARITY.	
FT	DISULFID 42 59	BY SIMILARITY.	
SQ	SEQUENCE 81 AA;	9087 MW;	0EEF7AB5CF6DC53 CRC64;

Query Match	45.68;	Score 154.5;	DB 5;	Length 81;	Best Local Similarity 51.7%;	Pred. No. 1.3e-12;	Matches 30; Conservative 7; Mismatches 18;	Indels 3;	Gaps 3;	Best Local Similarity 57.5%;	Pred. No. 1.3e-11;	Matches 23;	Conservative 6;	Mismatches 10;	Indels 1;	Gaps 1;			
Qy	5	ATIFIVLVAFCILEDGIVEAGFCGPNAKGKCHRCKSIRR-GGFCRGTFRTRTCVCY	61							QY	22	VEAGFGCPNAKGKCHRCKSIRR-RGGFCRGTFRTRTCVCY	60						
Db	3	AAFVLLVGCLIMTD-VATAGGCPNNYA-CHOHCKSTRGYGG'CAGNFRRLRCYR	58							Db	34	VRRGGCPENQYQCHSHESGIRGYKGCKTFFKQTCKY	73						
RESULT	4	Q9BLJ4	PRELIMINARY;	PRT;	73 AA.	Q9BLJ4	PRELIMINARY;	PRT;	73 AA.	RESULT	4	IAIIFIVLVAFCILE-DGI-----V	53	Query Match	41.38;	Score 140;			
RESULT	2	Q9U6U0	PRELIMINARY;	PRT;	82 AA.	Q9U6U0;	CDNA CLONING OF TICK DEFENSIN.;	PRT;	73 AA.	Best Local Similarity 40.3%;	Pred. No. 9e-11;	Matches 27;	Conservative 11;	Mismatches 19;	Indels 10;	Gaps 3;			
ID	Q9U6U0;	AC	Q9BLJ4;	AC	Q9BLJ4;	DT	01-JUN-2001 (TREMBLrel. 17, Created)	AC	Q9BLJ4;	AC	Q9BLJ4;	DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)	AC	Q9BLJ4;	AC			
DT	01-MAY-2000 (TREMBLrel. 13, Created)	DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)	DT	01-OCT-2001 (TREMBLrel. 18, Last annotation update)	DT	01-OCT-2001 (TREMBLrel. 18, Last annotation update)	DT	01-OCT-2001 (TREMBLrel. 18, Last annotation update)	DT	01-OCT-2001 (TREMBLrel. 18, Last annotation update)	DT	01-OCT-2001 (TREMBLrel. 18, Last annotation update)	DT	01-OCT-2001 (TREMBLrel. 18, Last annotation update)				
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)				
DE	ANTIMICROBIAL PEPTIDE MGDPB.	DE	ANTIMICROBIAL PEPTIDE MGDPB.	OS	Mytilus galloprovincialis (Mediterranean mussel).	OS	Eukaryota; Metazoa; Mollusca; Pteriomorphia; Mytiloida;	OC	Parasitiformes; Ixodida; Argasidae; Ornithodoros.										
OC	Mytiloidea; Mytilidae; Mytilus.	OX	NCBI_TAXID=29158;	RN	SEQUENCE FROM N.A.	RN	SEQUENCE FROM N.A.	RA	Nakajima Y.;	RA	Nakajima Y.;	RA	"CDNA CLONING OF TICK DEFENSIN.;"	RA	Mita G., Hubert F., Roch P.;				
RN	[1]	SEQUENCE FROM N.A.	RP	MITTA G., HUBERT F., ROCH P.;	RT	"Defensins: involvement in mussel defense.";													
RP	MITTA G., HUBERT F., ROCH P.;	RA	MITTA G., HUBERT F., ROCH P.;	RT	"Defensins: involvement in mussel defense.";	RT	"Defensins: involvement in mussel defense.";	RT	"Defensins: involvement in mussel defense.";	RT	"Defensins: involvement in mussel defense.";	RT	"Defensins: involvement in mussel defense.";	RT	"Defensins: involvement in mussel defense.";	RT	"Defensins: involvement in mussel defense.";		
RA	"Mytilus B and MG2, two antimicrobial peptides of marine mussels:	RT	"Mytilus B and MG2, two antimicrobial peptides of marine mussels:	RT	"Mytilus B and MG2, two antimicrobial peptides of marine mussels:	RT	"Mytilus B and MG2, two antimicrobial peptides of marine mussels:	RT	"Mytilus B and MG2, two antimicrobial peptides of marine mussels:	RT	"Mytilus B and MG2, two antimicrobial peptides of marine mussels:	RT	"Mytilus B and MG2, two antimicrobial peptides of marine mussels:	RT	"Mytilus B and MG2, two antimicrobial peptides of marine mussels:	RT	"Mytilus B and MG2, two antimicrobial peptides of marine mussels:		
RT	gene structure and expression analysis."	RT	gene structure and expression analysis."	RT	gene structure and expression analysis."	RT	gene structure and expression analysis."	RT	gene structure and expression analysis."	RT	gene structure and expression analysis."	RT	gene structure and expression analysis."	RT	gene structure and expression analysis."	RT	gene structure and expression analysis."		
RT	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.	RT	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.	RT	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.	RT	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.	RT	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.	RT	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.	RT	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.	RT	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.	RT	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.		
RL	AF017539; AADS2660.1;	DR	InterPro: IPR001542; Arthro_defensin.	DR	InterPro: IPR001542; Arthro_defensin.	DR	InterPro: IPR001542; Arthro_defensin.	DR	InterPro: IPR001542; Arthro_defensin.	DR	InterPro: IPR001542; Arthro_defensin.	DR	InterPro: IPR001542; Arthro_defensin.	DR	InterPro: IPR001542; Arthro_defensin.	DR	InterPro: IPR001542; Arthro_defensin.		
DR	InterPro: IPR001542; Arthro_defensin.	DR	InterPro: IPR001545; Glyco_hormone_beta.	DR	InterPro: IPR001545; Glyco_hormone_beta.	DR	InterPro: IPR001545; Glyco_hormone_beta.	DR	InterPro: IPR001545; Glyco_hormone_beta.	DR	InterPro: IPR001545; Glyco_hormone_beta.	DR	InterPro: IPR001545; Glyco_hormone_beta.	DR	InterPro: IPR001545; Glyco_hormone_beta.	DR	InterPro: IPR001545; Glyco_hormone_beta.		
DR	PFam: PF01097; Arthro_defensin.	DR	PROSITE: PS00425; ARTHROPOD_DEFENSINS; CRC64;	DR	PROSITE: PS00425; ARTHROPOD_DEFENSINS; UNKNOWN_1.	DR	PROSITE: PS00425; ARTHROPOD_DEFENSINS; UNKNOWN_1.	DR	PROSITE: PS00425; ARTHROPOD_DEFENSINS; UNKNOWN_1.	DR	PROSITE: PS00425; ARTHROPOD_DEFENSINS; UNKNOWN_1.	DR	PROSITE: PS00425; ARTHROPOD_DEFENSINS; UNKNOWN_1.	DR	PROSITE: PS00425; ARTHROPOD_DEFENSINS; UNKNOWN_1.	DR	PROSITE: PS00425; ARTHROPOD_DEFENSINS; UNKNOWN_1.		
DR	PFam: PF01097; Arthro_defensin.	DR	PS00261; GLYCO_HORMONE_BETA_1; UNKNOWN_1.	DR	PS00261; GLYCO_HORMONE_BETA_1; UNKNOWN_1.	DR	PS00261; GLYCO_HORMONE_BETA_1; UNKNOWN_1.	DR	PS00261; GLYCO_HORMONE_BETA_1; UNKNOWN_1.	DR	PS00261; GLYCO_HORMONE_BETA_1; UNKNOWN_1.	DR	PS00261; GLYCO_HORMONE_BETA_1; UNKNOWN_1.	DR	PS00261; GLYCO_HORMONE_BETA_1; UNKNOWN_1.	DR	PS00261; GLYCO_HORMONE_BETA_1; UNKNOWN_1.		
DR	SEQUENCE 82 AA;	DR	SEQUENCE 9307 MW;	DR	SEQUENCE 6D8F0039EF675ECE CRC64;	DR	SEQUENCE 6D8F0039EF675ECE CRC64;	DR	SEQUENCE 6D8F0039EF675ECE CRC64;	DR	SEQUENCE 6D8F0039EF675ECE CRC64;	DR	SEQUENCE 6D8F0039EF675ECE CRC64;	DR	SEQUENCE 6D8F0039EF675ECE CRC64;	DR	SEQUENCE 6D8F0039EF675ECE CRC64;	DR	SEQUENCE 6D8F0039EF675ECE CRC64;
DR	SEQUENCE 48.3%;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;
DR	Best Local Similarity 48.3%;	DR	Best Local Similarity 48.3%;	DR	Best Local Similarity 48.3%;	DR	Best Local Similarity 48.3%;	DR	Best Local Similarity 48.3%;	DR	Best Local Similarity 48.3%;	DR	Best Local Similarity 48.3%;	DR	Best Local Similarity 48.3%;	DR	Best Local Similarity 48.3%;	DR	Best Local Similarity 48.3%;
DR	Matches 28; Conservative 8;	DR	Matches 28; Conservative 8;	DR	Matches 28; Conservative 8;	DR	Matches 28; Conservative 8;	DR	Matches 28; Conservative 8;	DR	Matches 28; Conservative 8;	DR	Matches 28; Conservative 8;	DR	Matches 28; Conservative 8;	DR	Matches 28; Conservative 8;	DR	Matches 28; Conservative 8;
DR	Mismatches 20;	DR	Mismatches 20;	DR	Mismatches 20;	DR	Mismatches 20;	DR	Mismatches 20;	DR	Mismatches 20;	DR	Mismatches 20;	DR	Mismatches 20;	DR	Mismatches 20;	DR	Mismatches 20;
DR	Indels 2;	DR	Indels 2;	DR	Indels 2;	DR	Indels 2;	DR	Indels 2;	DR	Indels 2;	DR	Indels 2;	DR	Indels 2;	DR	Indels 2;	DR	Indels 2;
DR	Gaps 2;	DR	Gaps 2;	DR	Gaps 2;	DR	Gaps 2;	DR	Gaps 2;	DR	Gaps 2;	DR	Gaps 2;	DR	Gaps 2;	DR	Gaps 2;	DR	Gaps 2;
DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;
DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;
DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;
DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;
DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;
DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;
DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;
DR	Gaps 3;	DR	Gaps 3;	DR	Gaps 3;	DR	Gaps 3;	DR	Gaps 3;	DR	Gaps 3;	DR	Gaps 3;	DR	Gaps 3;	DR	Gaps 3;	DR	Gaps 3;
DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;
DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;
DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;
DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;
DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;
DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;
DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;
DR	Gaps 3;	DR	Indels 3;	DR	Gaps 3;	DR	Indels 3;	DR	Gaps 3;	DR	Indels 3;	DR	Gaps 3;	DR	Indels 3;	DR	Gaps 3;	DR	Indels 3;
DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;
DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;
DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;
DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;
DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;
DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;
DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;
DR	Gaps 3;	DR	Indels 3;	DR	Gaps 3;	DR	Indels 3;	DR	Gaps 3;	DR	Indels 3;	DR	Gaps 3;	DR	Indels 3;	DR	Gaps 3;	DR	Indels 3;
DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;	DR	Score 154.5;
DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;	DR	DB 5;
DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;	DR	Length 81;
DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;	DR	Best Local Similarity 51.7%;
DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;	DR	Pred. No. 1.3e-12;
DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;	DR	Matches 30;
DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;	DR	Indels 3;
DR	Gaps 3;	DR	Indels 3;	DR															

Best Local Similarity 28.8%; Pred. No. 0.0062; 21; Indels 29; Gaps 4;	DR InterPro; IPR002061; Scorpion_toxin.
Matches 23; Conservative 7; Mismatches 21; Indels 29; Gaps 4;	DR InterPro; IPR001219; Toxin.
QY 6 IIFIVLVAFCT-----LEDGIVE-----AGFGCPENAKGCHRCKS 41	DR Pfam; PF00537; Toxin_3; 1.
Db 5 IIVFAFIVAMCIAHSLAAPAPEALESVIIRQKRILCDLUSFEAKGFAA - NHSLCAAHCLA 62	DR PRINTS; PRO0084; TOXIN.
QY 42 IRRGGFCRTERFTCYCR 61	DR PRODOM; PD000908; Scorpion_toxin; 1.
Db 63 IGRGGACQ -- NGVCVCR 79	DR SMART; SM00505; Knot1; 1.
SEQUENCE 85 AA; 9588 MW; 416CB3D72A8927ED CRC64;	SQ
RESULT 9	Query Match 22.3%; Score 75.5; DB 5; Length 85;
ID Q9GYU6 PRELIMINARY; PRT; 57 AA.	Best Local Similarity 26.6%; Pred. No. 0.022;
AC Q9GYU6; 01-MAR-2001 (TREMBLrel. 16, Created)	Mismatches 13; Indels 7; Gaps 3;
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	Matches 17; Conservative 13; Mismatches 27; Indels 7; Gaps 3;
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)	QY 3 STALIFTWVAFCLIEDGIVEAGFGCPENAK--CHRHKSKIRRGFCR-GTERTC 57
DE DEFINSN (FRAGMENT).	Db 8 SLALLMTGVG--SVRDGYIADDRCNPYFCGRNAVCDGECKKRNRAESGICQWASKYGNAC 65
OS Aedes albopictus (Forest day mosquito).	QY 58 VCYR 61
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	Db 66 WCYK 69
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;	RESULT 11
OC	Q9GX2 PRELIMINARY; PRT; 85 AA.
OX NCBI_TAXID=7160;	NCBI_TAXID=34649;
RN [1]	Q9GX2 ID Q9GX2 PRELIMINARY; PRT; 85 AA.
RP SEQUENCE FROM N.A.	AC Q9GX2; 01-MAR-2001 (TREMBLrel. 16, Created)
RA Melinda P.M., Kostas B., Scott O.L.;	DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
RT "Wolbachia neither induces nor suppresses transcripts encoding	DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
RT antimicrobial peptides."; PREPRINTS; PRO0097; To the EMBL/GenBank/DBJ databases.	DE ALPHA_TOXIN PRECURSOR.
RL EMBL; AY005472; AAF97983.1; -.	OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
DR HSSP; P10891; 1ICA.	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones; Buthidae; Buthidae; Buthidae.
DR InterPro; IPR001542; Arthro_defensin.	NCBI_TAXID=34649;
DR Pfam; PF003614; Knot1.	RN [1]
DR PR00271; Arthro_defensin; 1.	RP SEQUENCE FROM N.A.
DR SMART; SM00505; Knot1; 1.	Ye J., Chen J., Zuo X., Ji Y.; Cloning and characterization of cDNA sequences encoding two novel alpha toxin precursors from the Chinese scorpion Buthus martensii.
DR PROSTIE; PS000425; ARTHROPOD_DEFENSINS; 1.	Karsch; Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
FT NON_TER 1 57	RL EMBL; AF288667; AA000580.1; -.
FT NON_TER 57 57	DR HSSP; P17728; IL01.
SQ SEQUENCE 57 AA; 6214 MW; A9F9DF5AE4CA86A7 CRC64;	DR InterPro; IPR002061; Knot1.
Query Match 22.7%; Score 77; DB 5; Length 57;	DR InterPro; IPR001219; Toxin.
Best Local Similarity 44.4%; Pred. No. 0.0095; 5; Mismatches 11; Indels 4; Gaps 2;	DR PRINTS; PRO0084; TOXIN.
Matches 16; Conservative 5; Mismatches 11; Indels 4; Gaps 2;	DR PRODOM; PD000908; Scorpion_toxin; 1.
QY 24 AGFGCPENAKGCHRCKSISTRRGFCRTERTCYC 59	DR SMART; SM00505; Knot1; 1.
Db 26 SGFGVGSAAHCIARRNGYCNAA - KTVVCV 57	KW SIGNAL.
SEQUENCE 85 AA; 9366 MW; 94ED463D613FFEF CRC64;	FT CHAIN 1 19 POTENTIAL.
RESULT 10	FT CHAIN 20 85 VCYR 61
Q9NJP8 PRELIMINARY; PRT; 85 AA.	SQ SEQUENCE 85 AA; 9366 MW; 94ED463D613FFEF CRC64;
ID Q9NJP8	Query Match 22.3%; Score 75.5; DB 5; Length 85;
AC Q9NJP8; 01-OCT-2000 (TREMBLrel. 15, Created)	Best Local Similarity 26.6%; Pred. No. 0.022;
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)	Mismatches 12; Indels 7; Gaps 3;
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)	QY 3 STALIFTWVAFCLIEDGIVEAGFGCPENAK--CHRHKSKIRRGFCR-GTERTC 57
DE NEUROTOXIN TX11P.	Db 8 SLALLMTGVG--SVRDGYIADDRCNPYFCGRNAVCDGECKKRNRAESGICQWASKYGNAC 65
OS Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones; Buthidae; Buthidae; Buthidae.	QY 58 VCYR 61
OC	Db 66 WCYK 69
RN SEQUENCE FROM N.A.	RESULT 12
RC TISSUE=VENOM GLAND;	Q9BLM4 PRELIMINARY; PRT; 85 AA.
RA Zhu S.; Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases	ID Q9BLM4
DR EMBL; AF32974; AAT31205.1; -.	AC Q9BLM4; 01-JUN-2001 (TREMBLrel. 17, Created)
HSSP; P01484; 1AO.	DR
InterPro; IPR003614; Knot1.	DR

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE PUTATIVE TOXIN PRECURSOR.
 DE Androctonus australis (Sahara scorpion).
 OS Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Buthidae; Buthidae; Androctonus .
 NCBI_TAXID=6888;

[1] RN
 SEQUENCE FROM N.A.
 RX MEDLINE=21211144; PubMed=1131124;
 RA Céard B., Martin-Bauclaire M.F., Bougis P.E.;
 RT "Evidence for a position-specific deletion as an evolutionary link
 between long- and short-chain scorpion toxins.";
 RL PEBS Lett. 49(4):246-248(2001).
 DR EMBL: AJ084440; HSSP: P01484; IAH0;
 DR InterPro: IPRO03614; Knot1.
 DR InterPro: IPRO02061; Scorpion_toxin.
 DR InterPro: IPRO01219; Toxin.
 DR ProtDom: PD00008; Scorpion_toxin; 1.
 DR SMART: SM00505; Knot1; 1.
 KW SIGNAL.
 FT CHAIN 1 19 POTENTIAL.
 ST SEQUENCE 85 AA; 9523 MW; 4059A69D0E4F090 CRC64;

Query Match 21.1%; Score 71.5%; DB 5; Length 85;
 Best Local Similarity 27.3%; Pred. No. 0.072;
 Matches 18; Conservative 12; Mismatches 29; Indels 7; Gaps 3;

Qy 1 MKSIAIIIFVLAQICILEGGIVEAGFGCFNACK---CHRCKSIRRGFCR--GTFRT 55
 . Db 6 MISALLFTWGV--SKKGYIVDDKNCTFFCGRNAYCNDEEKKKGAESGYCOWASPYGN 63

Qy 56 TCVCYR 61
 . Db 64 ACYCYK 69

RESULT 13
 ID 061721 PRELIMINARY; PRT; 102 AA.
 AC 061721; PRT; 102 AA.
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DE DEFENSIN.
 GN DEF.
 OS Anopheles gambiae (African malaria mosquito).
 OC Pterygota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Anophelles.
 NCBI_TAXID=7165;

RN
 RP SEQUENCE FROM N.A.
 RC Eggleston P., Lu W., Zhao Y.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR HSSP: AF063402; AAC18575.1; -.
 DR InterPro: IPR01542; Arthro_defensin.
 DR Pfam: PF01097; Arthro_defensin; 1.
 DR PRINTS: PR00271; DEFENSIN.
 DR SMART: SM00505; Knot1; 1.
 DR PROSITE: PS00425; ARTHROP_DEFENSINS; 1.
 SQ SEQUENCE 1061 AA; 6288289416200DC CRC64;

Query Match 20.6%; Score 70; DB 5; Length 102;
 Best Local Similarity 41.7%; Pred. No. 0.13;

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE PUTATIVE TOXIN PRECURSOR.
 DE Androctonus australis (Sahara scorpion).
 OS Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Buthidae; Buthidae; Androctonus .
 NCBI_TAXID=6888;

[1] RN
 SEQUENCE FROM N.A.
 RX MEDLINE=21211144; PubMed=1131124;
 RA Céard B., Martin-Bauclaire M.F., Bougis P.E.;
 RT "Evidence for a position-specific deletion as an evolutionary link
 between long- and short-chain scorpion toxins.";
 RL PEBS Lett. 49(4):246-248(2001).
 DR EMBL: AJ084440; HSSP: P01484; IAH0;
 DR InterPro: IPRO03614; Knot1.
 DR InterPro: IPRO01219; Toxin.
 DR ProtDom: PD00008; Scorpion_toxin; 1.
 DR SMART: SM00505; Knot1; 1.
 KW SIGNAL.
 FT CHAIN 1 19 POTENTIAL.
 ST SEQUENCE 85 AA; 9523 MW; 4059A69D0E4F090 CRC64;

Query Match 21.1%; Score 71.5%; DB 5; Length 85;
 Best Local Similarity 27.3%; Pred. No. 0.072;
 Matches 18; Conservative 12; Mismatches 29; Indels 7; Gaps 3;

Qy 1 MKSIAIIIFVLAQICILEGGIVEAGFGCFNACK---CHRCKSIRRGFCR--GTFRT 55
 . Db 6 MISALLFTWGV--SKKGYIVDDKNCTFFCGRNAYCNDEEKKKGAESGYCOWASPYGN 63

Qy 56 TCVCYR 61
 . Db 64 ACYCYK 69

RESULT 15
 ID 09FR81 PRELIMINARY; PRT; 72 AA.
 AC 09FR81; PRT; 72 AA.
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE DISEASE RESISTANCE RESPONSE PROTEIN 230.
 GN DRR230.

OS Pisum sativum (Garden pea).
 OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Rosidae;
 OC NCBI_TAXID=3888;

RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV_GREENFEAST;
 RA Savenstrand H., Brosche M., Strid A.;
 RT "Stress-induced disease resistance protein 230 CDNA from
 RT Pisum sativum cv. Greenfeast."
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF139018; AAG43285.1; -.
 DR InterPro: IPR002118; Gamma-thionin.
 DR InterPro: IPR003614; Knot1.
 DR Pfam: PF00304; Gamma-thionin; 1.
 DR ProDom: PD002594; Gamma-thionin; 1.
 DR SMART: SM00305; Knot1; 1.
 SQ SEQUENCE 72 AA; 8000 MW; 00F9B605FED920 CRC64;

Query	Match	20.5%	Score	69.5	DB	10	Length	72
Best local Matches	20;	Similarity Conservative	29.9% 10;	Pred. Mismatches	No. 0.11; <td></td> <td></td> <td></td>			
						24:	Indels	13;
							Gaps	3;
Qy	2	KSTIAIIFIVLVAFACILED GIVEAGFGCPFNAGK -----						
Ddb	4	KSLACLSELFLVLFIAQETIVSEAN <u>T</u> CENLAGS <u>SYR</u> CVCGCDRHICRT--QEGATSGR <u>C</u> R						
Qy	51	GFRFTTC	57					
Ddb	62	D ₁ RCW ₁	68					

Search completed: September 17, 2002, 15:45:33
Job time: 129 sec